University of Alberta

Genetic Mapping of Disease Resistance Homologues in *Brassica napus* and *Arabidopsis thaliana*: Identification of Candidate Resistance Genes

by

Lee Dean Sillito

©

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for
the degree of Master of Science

Department of Biological Sciences

Edmonton, Alberta

Spring 2000

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-60176-5

Canada

University of Alberta

Library Release Form

Name of Author: Lee Dean Sillito

Title of Thesis: Genetic Mapping of Disease Resistance Homologues in *Brassica napus* and *Arabidopsis thaliana*: Identification of Candidate Resistance Genes

Degree: Master of Science

Year this Degree Granted: 2000

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Dean Sillito

Lee Dean Sillito
10983 - 72nd Ave. NW.
Edmonton, Alberta, Canada
T6G 0B1

Date: Dec. 20, 1999

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "Genetic Mapping of Disease Resistance Homologues in *Brassica napus* and *Arabidopsis thaliana*: Identification of Candidate Resistance Genes" submitted by Lee Dean Sillito in partial fulfillment of the requirements for the degree of Master of Science

Allen Good

Dr. Allen Good
(Supervisor)

Dr. Peter Constabel

Dr. Ross Hodgetts

Dr. Francis Yeh
(External Examiner)

Abstract

Common structural motifs among cloned plant disease resistance genes (R genes) have made possible the identification of putative disease resistance sequences based on DNA sequence homology. Mapping these R gene homologues will identify candidate disease resistance loci to expedite map based cloning strategies in complex crop genomes. *Arabidopsis thaliana* expressed sequence tags (ESTs) with homology to cloned plant R genes (R-ESTs) were mapped in *A. thaliana* and *Brassica napus* to identify candidate R gene loci and investigate intergenomic collinearity. *Brassica* R gene homologues were also mapped in *B. napus*. In total, 103 R-EST loci and 36 *Brassica* R gene homologues were positioned on the *B. napus* map, and 48 R-EST loci were positioned on the *A. thaliana* map. Mapped loci identified collinear regions in *Brassica* for *Arabidopsis* chromosomes 1, 3, 4, and 5. Detection of syntenic regions indicated there was no apparent rapid divergence of genomic regions housing the R-ESTs.

Acknowledgements

I would like to thank Dr. Allen Good for supervising and providing support for this research, as well as my committee members Dr. Peter Constabel, Dr. Ross Hodgetts and Dr. Gary Stringam.

Thanks go to Dr. Derek Lydiate's group at Agriculture Canada in Saskatoon for allowing me to work in their labs, and allowing access to the *B. napus* mapping population financially supported by Zeneca Seeds.

Thanks to everyone that I worked with in the lab at the University of Alberta, as well as in Saskatoon for interesting discussions, making me feel welcome, and showing me a good time. I would also, in particular, like to thank Dr. Reinhold Mayerhofer for teaching me the skills for genetic mapping, and Dr. Isobel Parkin for collaboration on paper-writing and continuously reminding me how little I really know!

Financial support for this research was provided by the Natural Sciences and Engineering Research Council, Agriculture Canada, the Alberta Agricultural Research Institute, and the Saskatchewan Agri-Food Innovation Fund.

Table of Contents

I. Introduction 1

1.1 Impact of Pathogens on Agriculture 1

1.2 Breeding Disease Resistance with Modern Biotechnology 2

1.3 The Gene-for-Gene Nature of Plant-Pathogen Interactions 4

1.4 Signal Cascades Lead to Plant Defense Responses 6

1.4.1 Reactive Oxygen Species are Formed Subsequent to Pathogen Attack 6

1.4.2 Identification of Signal Transduction Genes 8

1.4.3 Disease Recognition Leads to Synthesis of Protective Compounds and Induction of SAR 10

1.5 R Genes are Classified According to Common Structural Motifs 11

1.5.1 Common R Gene Structural Motifs 11

1.5.2 R Gene Classes 16

1.6 Comparative Genetic Mapping 16

1.7 Exploitation of R Gene Homology in Conserved Motifs 19

1.7.1 PCR Amplification of R Gene Homologues Using Degenerate Primers 19

1.7.2 Utilizing Expressed Sequence Tags in the R Gene Candidate Search 22

1.8 Evolution of R Gene Loci 24

1.9 *Brassica* Genome Relationships 29

1.10 Mapped *Brassica* R Genes 31

1.11 Objectives of This Research 32

II. Materials Methods 33

2.1 Plant Mapping Populations 33

2.2 Genomic DNA Extraction and Southern Hybridization 35

2.3 RFLP Probes 36

2.4 PCR Conditions 37

2.5 Linkage Analysis 37

2.6 Brief Overview of RFLP Map Generation 38

2.7 Atypical RFLP Patterns 40

III. Results and Discussion 42

3.1 Genetic Mapping of *Arabidopsis* R-ESTs 42

3.1.1 R-EST Loci Detected in A. thaliana and B. napus 42

3.1.2 Duplication of R-EST Loci in A. thaliana 50

3.1.3 Loci Organized as Tandem Repeats 51

3.2 Genetic Mapping of *Brassica* R Gene Homologues 53

3.3 Collinearity Between *Arabidopsis* and *Brassica* Genomes Detected by R-EST Loci 56

3.5 Conclusion 64

IV. References 67

List of Tables

Table 1 Plant disease resistance genes cloned to date 12-13
Table 2 *Arabidopsis* R-EST loci mapped in *B. napus* and *A. thaliana* 43-44
Table 3 *B. napus* disease resistance homologue loci mapped in *B. napus* 54

List of Figures

Figure 1 Receptor-ligand model of gene-for-gene resistance 5
Figure 2 Components of the signaling cascade leading to resistance 7
Figure 3 U's triangle showing genomic relationship of Brassicas 30
Figure 4 Generation of N-fo-61-9 *B. napus* mapping population 34
Figure 5 Scoring autoradiographs and map construction 39
Figure 6 Typical RFLP banding pattern seen with disomic inheritance 41
Figure 7 *B. napus* linkage map with R-ESTs and *Brassica* homologues 46-47
Figure 8 *A. thaliana* linkage map with R-ESTs 48-49
Figure 9 Multiple co-segregating RFLP bands indicative of tandem repeats 52
Figure 10 *B. napus* collinearity with *A. thaliana* chromosome 3 58
Figure 11 *B. napus* collinearity with *A. thaliana* chromosome 4 60
Figure 12 *B. napus* collinearity with *A. thaliana* chromosome 1 62
Figure 13 Overview of collinearity between the A and C genomes of *B. napus* 63

Abbreviations

Avr	-	Avirulence
BAC	-	Bacterial Artificial Chromosome
CAPS	-	Cleavable Amplified Polymorphic Sequence
cDNA	-	Complementary Deoxyribonucleic Acid
cm	-	Centimetre
cM	-	Centimorgan
Col-0	-	Columbia accession of *A. thaliana*
cv.	-	Cultivar
dATP	-	Deoxyadenine Triphosphate
dbEST	-	Expressed Sequence Tag Database
dCTP	-	Deoxycytosine Triphosphate
dGTP	-	Deoxyguanine Triphosphate
DH	-	Doubled Haploid
DNA	-	Deoxyribonucleic Acid
dTTP	-	Deoxythimine Triphosphate
EDTA	-	Ethylenediamine Tetraacetic Acid
EST	-	Expressed Sequence Tag
F1	-	First Filial Generation
g	-	Gram
HC	-	Highly Conserved
HR	-	Hypersensitive Response
hrp	-	Hypersensitive Response and Pathogenesis
Ka	-	Non-synonymous amino acid substitution
kb	-	Kilobase Pairs
Ks	-	Synonymous amino acid substitution
Ler-0	-	Landsberg *erecta* accession of *A. thaliana*
LG	-	Linkage Group
LOD	-	Logarithm of the Odds
LRR	-	Leucine Rich Repeat
LZ	-	Leucine Zipper
Mbp	-	Megabase pairs
min	-	Minutes
mM	-	Millimolar
MRC	-	Major Recognition Complex
mRNA	-	Messenger Ribonucleic Acid
n	-	Haploid Chromosome Complement
NBS	-	Nucleotide Binding Site
ng	-	Nanogram
PCR	-	Polymerase Chain Reaction
PR	-	Pathogenesis Related
QTL	-	Quantitative Trait Loci
R gene	-	Plant Disease Resistance Gene
R-EST	-	Resistance Expressed Sequence Tag

Abbreviations (continued)

RFLP	-	Restriction Fragment Length Polymorphism
RGA	-	Resistance Gene Analog
RI	-	Recombinant Inbred
RNAse	-	Ribonuclease A
ROS	-	Reactive Oxygen Species
SA	-	Salicylic Acid
SAR	-	Systemic Acquired Resistance
SDS	-	Sodium Dodecyl Sulfate
Ser/Thr	-	Serine/Threonine
SSC	-	Salt Sodium Citrate
TAE	-	Tris Acetate EDTA
TIR	-	Toll/Interleukin-1 Receptor Homologous Domain
TM	-	Transmembrane
μl	-	Microlitre
μM	-	Micromolar
μg	-	Microgram
V	-	Volts
YAC	-	Yeast Artificial Chromosome

I. Introduction

1.1 Impact of Pathogens on Agriculture

All crop plants grown for human consumption are attacked and consumed by a myriad of pathogenic organisms. Plant diseases may be caused by bacteria, viruses, fungi or nematodes, and in recent history there has been great progress in understanding the dynamics of these plant-pathogen interactions, as well as developing new methods to minimize disease-induced crop loss.

It is important to study plant diseases, as they can have severe impacts on human activities and survival. Severe crop epidemics such as the Irish potato famine of the mid 1840's caused by an outbreak of the fungus *Phytophthora infestans* which resulted in the loss of one quarter of Ireland's population, including 1 million starvation deaths, have shown how dependent we are on agriculture, and how crop diseases can rapidly affect our lives (Agrios 1988; Strange 1993). In the industrialized world, disease losses are most often seen in economic terms, as populations are less directly tied to the output of the land than subsistence farmers. Agricultural output is high and diversified for both internal use and export, due to mechanization and advanced farming technology. It has been estimated that over $100 billion are lost on a yearly basis in the world due to crop disease (Brears and Ryals 1994). Crop yield increases due to improved cultivars, fertilization, and cultivation techniques must be protected from disease loss to ensure a constant and reliable food supply (Agrios 1988).

A number of disease control methods are available to the agricultural industry. Cultural practices such as crop rotation, and appropriate irrigation and fertilization are some of the

oldest disease control methods, and are still effective today (Kaminski *et al.* 1997). Application of chemicals such as fungicides and insecticides (to control insect disease vectors) has been the mainstay of modern agricultural disease control, but the chemicals used are generally toxic to non-target organisms in the environment, and if persistent, can leach into groundwater supplies, rendering them unusable. Although highly effective, the toxicity and expense of intensive chemical usage has prompted the development of more environmentally sound disease control methods which utilize the natural sources of genetic resistance found within plants themselves (Yang *et al.* 1997).

1.2 Breeding Disease Resistance with Modern Biotechnology

The creation of disease resistant cultivars by plant breeders has been successful in reducing disease losses and chemical applications. Breeding of disease resistance generally involves crossing a resistant plant to an agronomically superior susceptible plant, followed by backcrossing to the susceptible parent (recurrent parent) in an attempt to recover as much of the superior parent's genetic makeup as possible, while still retaining the resistance phenotype introduced from the non-recurrent parent. Sources of resistance are sought from existing cultivars or breeding lines, or from wild landraces, usually found growing in the centre of diversity where the crop plant in question originated (Agrios 1988). Traditional breeding methods come up against a natural barrier in the necessity for the two parents of a cross to be sexually compatible; usually of the same, or highly related species.

Biotechnological advances are being incorporated by breeders as they become available

and demonstrate their utility. Molecular markers linked to desirable traits, improved methods of product analysis, and embryo and tissue culture techniques are all being used to improve the speed, efficiency, and expand the boundaries of plant breeding (Woolhouse 1992). Currently, biotechnological methods can be used to physically transfer genes conferring disease resistance between widely related species that cannot form natural hybrids. In order for disease resistance engineering strategies to be successful, the mechanisms of plant resistance at the molecular level should be understood.

Many genes in a given plant are useful for pathogen defense, including those responsible for preformed barriers to infection such as the cuticle and cell wall components. Genes which result in resistance to specific races of a particular pathogen have been most intensely studied in recent years. These have been designated as R (resistance) genes (Strange 1993) and show a characteristic gene-for-gene interaction between pathogen and host.

Detailed and complete genetic maps will allow plant breeders to better identify, manipulate, and combine traits for maximum utility (Allen 1994). Precise maps will help provide DNA markers in chromosomal regions of interest. They will also help distinguish tightly linked genes from genes with pleiotropic effects, as it is desirable to know if associations between traits can be broken by recombination (Allen 1994). Indirect selection in a breeding program using linked markers can be beneficial if phenotypic selection is inaccurate, costly, or time consuming. DNA markers can identify qualitative genes or members of qualitative trait loci (QTL), help determine effects of gene dosage,

and give information on the chromosomal locations of each gene affecting a particular trait.

The cloning of disease resistance genes and homologues will assist in breeding disease resistance by providing perfect makers which are unlikely to be separated by recombination events (Spielmeyer *et al.* 1998), and R genes can be tracked through a breeding program without the need for progeny testing by pathogen inoculation, which often requires plants to be grown to maturity (Staskawicz *et al.* 1995). Markers linked to R genes will assist in the pyramiding of R genes, which is a strategy thought to increase the durability of resistance in the field.

1.3 The Gene-for-Gene Nature of Plant-Pathogen Interactions

Insights into how R genes function first came from H. Flor (1956) in his work studying the interaction between flax (*Linum ultissimum*) and flax rust (*Melampsora lini*). After observing the disease response from various combinations of host and pathogen genotypes, Flor proposed that compatibility (virulence/susceptibility) or incompatibility (avirulence/resistance) was determined by the interaction of a gene in the host and a gene in the pathogen. Race specific resistance depends on a dominant R gene in the plant and its interaction with a dominant Avr (avirulence) gene in the pathogen, resistance only occurring with a dominant gene at each locus (Figure 1) (Hammond-Kosack and Jones 1995). A loss or mutation in either the dominant R gene or its corresponding Avr gene would result in a compatible reaction, and disease would ensue. Only in the past decade

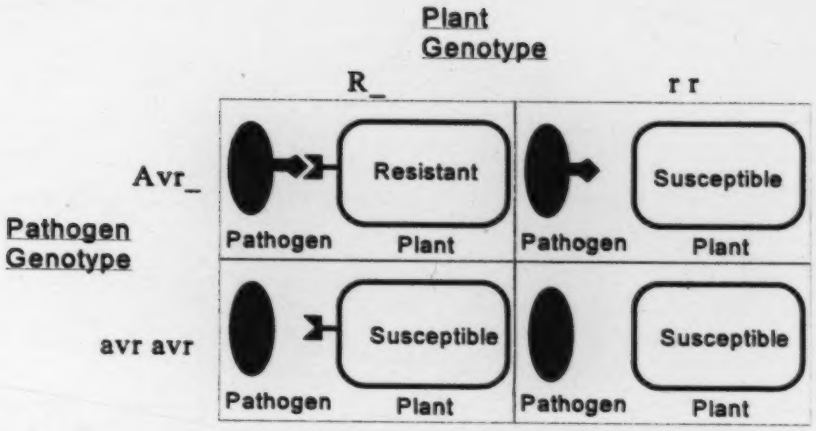


Figure 1. Possible interactions between a plant R gene and a pathogen Avr gene according to the gene-for-gene model of disease resistance. Only when alleles at both corresponding loci are dominant will recognition of the pathogen take place, initiating signal cascades leading to defense responses.

have R and Avr genes been cloned in order to study their interaction at the molecular level. Avr genes expressed by pathogens are thought to have alternate survival or pathogenicity function, and are not simply determinants for their own recognition by R genes and subsequent destruction (Dangl 1994). Host plants have evolved to recognize the products of Avr genes and trigger a cascade of events resulting in the ultimate destruction of the pathogen and resistance for the plant.

In the current model of Avr/R gene interaction, the product of the Avr gene, whatever its function, becomes an elicitor recognized by an intra- or extracellular receptor encoded by the R gene. Intracellular receptors would initiate the disease response mechanism by binding viral elicitors or virulence components injected into the host cells via the

bacterial hrp type III protein secretion pathway (Yang *et al.* 1997). The receptors would then interact with other components in a signaling cascade ultimately leading to disease resistance.

Support for the gene-for-gene theory has accumulated for many R/Avr gene combinations. For example, pathogen races transgenic for a particular Avr gene not naturally expressed will adopt the virulence characteristics of the pathogen from which the Avr gene was introduced (Staskawicz *et al.* 1984; Thilmony *et al.* 1995). The tomato *Pto* gene conferring resistance to *P. syringae* pv. tomato was shown to bind directly to the product of the AvrPto fungal gene in a yeast two hybrid system (Tang *et al.* 1996), implying a receptor-type function.

1.4 Signal Cascades Lead to Plant Defense Responses

Once an incompatible R/Avr gene reaction has taken place, a host reaction called the hypersensitive response (HR) occurs, where tissue in the area surrounding the site of infection undergoes rapid necrosis to prevent spread of the pathogen. This hypersensitive response can also trigger a secondary, nonspecific pathogen protection cascade, systemic acquired resistance (SAR) which can last for several days (Baker *et al.* 1997) (Figure 2).

1.4.1 Reactive Oxygen Species are Formed Subsequent to Pathogen Attack

The production of reactive oxygen species (ROS) is one of the first observed events following pathogen recognition and plays a key role in plant defense and induction of the HR (Baker and Orlandi 1995; Hammond-Kosack and Jones 1996; Tenhaken *et al.* 1995).

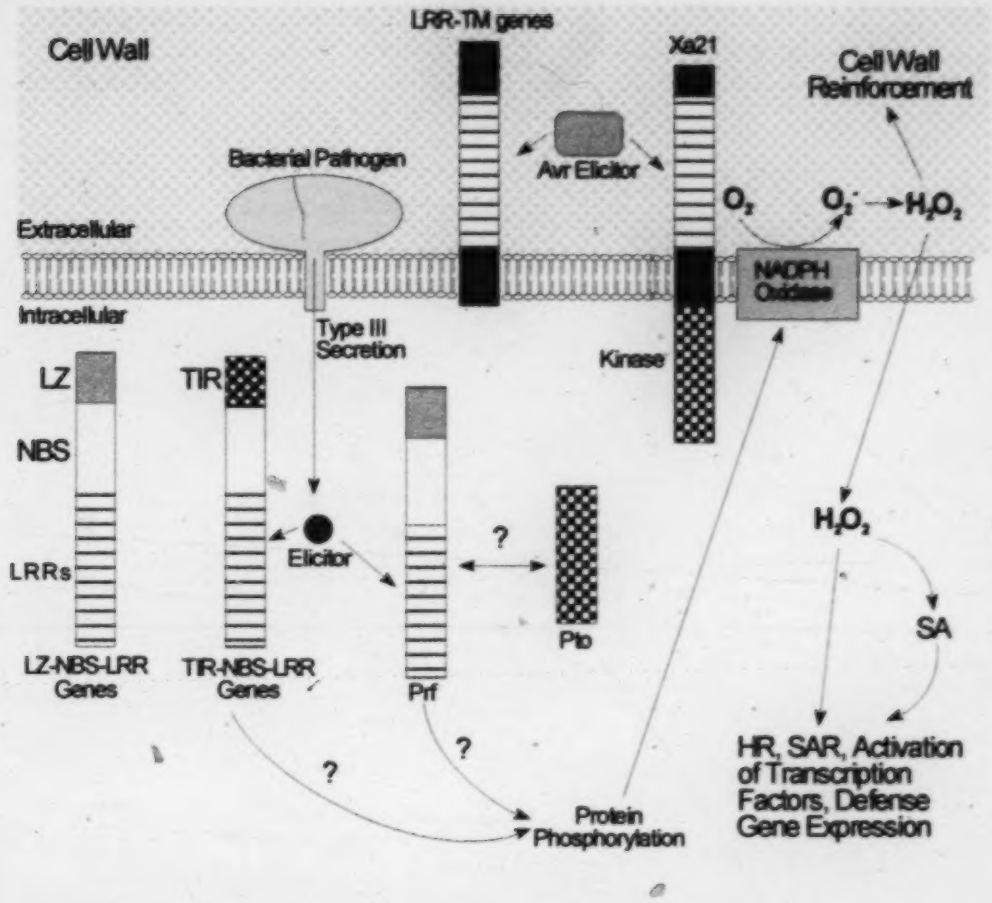


Figure 2. Simplified diagrammatic representation of the proposed locations of various R genes, and some of the interactions that occur with Avr gene-encoded elicitors. R genes are depicted with their common amino acid motifs, and are oriented with their N-terminal regions towards the top of the page, and the C-terminal regions toward the bottom. Pto can directly bind AvrPto, and other genes are thought to bind their corresponding Avr products either directly, or coupled with another binding protein. Prf is required for Pto function, although its role is unknown. Protein phosphorylation events brought about by ligand recognition are believed to activate NADPH oxidase and stimulate the production of hydrogen peroxide (see text) (Hammond-Kosack and Jones, 1997; Yang et al., 1997). LRR- leucine rich repeat; NBS- nucleotide binding site; TIR- Toll/Interleukin-1 Receptor domain; LZ- leucine zipper; TM- transmembrane domain; SA- salicylic acid; SAR- systemic acquired resistance; HR- hypersensitive response.

The hydroperoxyl radical $HO_2\cdot$ is formed and can directly attack lipid molecules, leading to membrane damage and formation of lipid peroxide signal molecules (Hammond-Kosack and Jones 1996). An increase in hydrogen peroxide (H_2O_2) concentration is thought to have multiple functions in resistance. In the presence of Fe^{2+}, H_2O_2 forms $OH\cdot$ (hydroxyl free radical), which can degrade DNA through site specific attack. ROS could be directly toxic to microbes, and can increase the strength of cell walls, as H_2O_2 is required for the formation of lignin precursors (Levine *et al.* 1994). Cell wall fortification is a common result of pathogen recognition, thereby impeding the pathogen-induced breakdown of the cell wall to gain entry to the cell, or to release nutrients. H_2O_2 could also be involved in salicylic acid (SA) biosynthesis, which is used as a signal molecule to induce the production of pathogen related genes in some plant species, and could also be involved in increasing the efficacy of ROS produced by inhibiting their breakdown by catalases (Ryals *et al.* 1996).

1.4.2 Identification of Signal Transduction Genes

Disease resistance mutants have been particularly useful in delineating the signal transduction pathway in *A. thaliana* (Century *et al.* 1997; Falk *et al.* 1999; Frye and Innes 1998). More R genes than signal transduction components are identified in these studies however, suggesting that transduction components have built in redundancy mechanisms or are required for viability (Warren *et al.* 1999).

Mutagenesis studies in *A. thaliana* have identified several genes required in the signal transduction pathway downstream of pathogen recognition by R genes. The *NDR1* gene is

required for the function of several disease resistance genes conferring resistance against pathogens including *P. syringae* pv tomato expressing avrB, avrRp2, avrRpm1 or avrPph3, and numerous *P. parasitica* isolates, making it a strong candidate for a conserved signal transduction element downstream of R genes with diverse specificity (Century *et al.* 1997). Complementation of various *ndr1* mutants also supports a vital role for the *NDR1* protein in mobilizing disease resistance. The structure of *NDR1* contains two transmembrane segments, so it is believed to be membrane-associated.

EDS1 is another *Arabidopsis* gene required for R gene function. It has blocks of conserved amino acids homologous to the catalytic domain of eukaryotic lipases (Falk *et al.* 1999). It is required upstream of SA in the signaling pathway, and is thought to hydrolyze a lipid, or other substrate to generate further signaling molecules.

Different subsets of *A. thaliana* R genes are dependent on either *NDR1* or *EDS1*, implying the presence of at least two distinct signaling pathways following R gene recognition. Function of *RPP2*, *RPP4*, *RPP5*, *RPP21* and *RPS4* expressing avrRps4 are all dependent on the presence of *EDS1*, whereas *RPS2*, *RPM1* and *RPS5* are dependent on *NDR1* (Aarts *et al.* 1998b). It appears to be R gene structure rather than any pathogen characteristic which determines the favoured pathway. *EDS1* associated genes are of the TIR-NBS-LRR type, and *NDR1* associated genes have an LZ-NBS-LRR structure. There does appear to be a slight amount of 'cross-talk' between the two pathways however, as *EDS1* dependent genes *RPP4* and *RPP5* show a slight but significant disease response reduction in *ndr1* mutant plants (Aarts *et al.* 1998b).

It has recently been discovered that *RPM1* (an *NDR1*-dependent R gene) despite having

no recognizable transmembrane motif, is a peripheral plasma membrane protein likely residing on the cytoplasmic membrane face (Boyes *et al.* 1998). Since NDR1 has transmembrane domains, it is believed that RPM1 must complex with NDR1 to conduct the recognition signal into the cell. A similar gene relationship has been observed between *RPS5* and *PBS1*; both required for the recognition of avrPphB. Linkage between these two genes also suggests a functional association, which has been observed with other genes such as *Pto* and *Prf*, and the *Brassica S* locus genes (Warren *et al.* 1999). Whether the *NDR1/RPM1* type of interaction is occurring with Pto/Prf and *RPS5/PBS1* remains to be determined. It is possible that other NBS-LRR cytosolic R genes may require similar associations with unidentified membrane bound proteins.

1.4.3 Disease Recognition Leads to Synthesis of Protective Compounds and Induction of SAR

Several pathogen related (PR) proteins are induced by the signal cascade resulting from pathogen attack (Figure 2). These proteins include glucanases and chitinases which degrade bacterial and fungal cell walls respectively, lipoxygenases which could attack both pathogen and host cells, and phytoalexins, which are a group of lipophilic antimicrobial compounds which have been observed to accumulate at the site of infection (Baker and Orlandi 1995).

Systemic acquired resistance is the response of a plant to pathogen attack which serves to increase the disease resistance capacity of the entire plant for a sustained period. This particular signal cascade is induced by cell death resulting from the HR or by a

compatible disease reaction (Ryals *et al.* 1996). It is distinguished by the broad spectrum of pathogen resistance characteristics, and the selective expression of certain defense related genes (SAR genes). SAR has been shown to depend on increased levels of salicylic acid within the plant. SA is not believed to be the translocated signal to induce SAR in plant regions distal to the site of infection, and is thought to be downstream of the unknown transducer, although this has not been conclusively demonstrated. Experimental inhibition of SAR compromises plant resistance, and stimulation of SAR increases resistance (Ryals *et al.* 1996).

1.5 R Genes are Classified According to Common Structural Motifs

1.5.1 Common R Gene Structural Motifs

Cloned R genes can be grouped into categories based on common DNA and protein motifs (Table 1). The realization that R genes from a widely diverse group of plants, and against all types of pathogens have structural homology implies a conserved evolutionary mode of function. It is these common motifs which are the basis of isolating R gene homologues, which may turn out to have resistance function, and be used as candidate R genes in map-based cloning strategies.

The most common motif found among cloned and characterized plant R genes is the leucine rich repeat (LRR). These eukaryotic conserved sequences encode LRR proteins found both cytoplasmically and extracytoplasmically. The general amino acid consensus sequence for extracytoplasmic LRRs is LXXLXXLXXLXLXXNXLXGXIPXX, and for cytoplasmic LRRs is LXXLXXLXXLXLXX(N/CK)X(X)LXXIPXX (Jones and Jones

[a] Class designations according to Baker et al. (1997).

[b] Structural motifs; LRR-leucine rich repeat; NBS- nucleotide binding site; LZ- leucine zipper; TIR- Toll / Interleukin-1 Receptor, TM- transmembrane domain

[c] References: (1) (Johal and Briggs 1992); (2) (Martin *et al.* 1993); (3) (Bent *et al.* 1994);(4) (Mindrinos *et al.* 1994);(5) (Grant *et al.* 1995);(6) (Ori *et al.* 1997); (7) (Whitham *et al.* 1994); (8) (Lawrence *et al.* 1995); (9) (Anderson *et al.* 1997);(10) (Botella *et al.* 1998); (11) (Parker *et al.* 1997);(12) (Dixon *et al.* 1996);(13) (Thomas *et al.* 1997);(14) (Dixon *et al.* 1998);(15) (Jones *et al.* 1994);(16) (Cai *et al.* 1997);(17) (Song *et al.* 1995);(18) (Salmeron *et al.* 1996);(19) (Warren *et al.* 1998);(20) (McDowell *et al.* 1998);(21) (Lagudah *et al.* 1997).

Table 1. Cloned plant R gene classes showing source plants, pathogens, and common structural motifs.

Class[a]	R Gene	Plant	Pathogen	Avr Gene	R Gene Structure[b]	Ref.[c]
1	RPS2	*Arabidopsis*	*Pseudomonas syringae* pv. tomato	avrRpt2	LZ-NBS-LRR	(3,4)
	RPM1	*Arabidopsis*	*Pseudomonas syringae* pv. maulicola	avrRpm1, avrB	LZ-NBS-LRR	(5)
	Prf	tomato	*Pseudomonas syringae* pv. tomato	avrPto	LZ-NBS LRR	(18)
	RPS5	*Arabidopsis*	*Pseudomonas syringae* pv. tomato DC3000	avrPphB	LZ-NBS-LRR	(19)
	RPP8	*Arabidopsis*	*Peronospora parasitica*	unknown	LZ-NBS-LRR	(20)
	Mi	tomato	*Meloidogyne* spp.	unknown	LZ-NBS-LRR	(22)
	N	tobacco	mosaic virus	unknown	TIR-NBS-LRR	(7)
	L^6	flax	*Melampsora lini*	Al^6	TIR-NBS-LRR	(8)
	M	flax	*Melampsora lini*	AM	TIR-NBS-LRR	(9)
	RPP1	*Arabidopsis*	*Peronospora parasitica*	unknown	TIR-NBS-LRR	(10)
	I_2	tomato	*Fusarium oxysporum* f. sp. lycopersicon	unknown	NBS-LRR	(6)
	Cre3	wheat	*Heterodera avenae*	unknown	NBS-LRR	(21)
	RPP5	*Arabidopsis*	*Peronospora parasitica*	avrPp5	TIR-NBS-LRR	(11)
2	Pto	tomato	*Pseudomonas syringae* pv. tomato	avrPto	ser/thr protein kinase	(2)
3	Cf-2	tomato	*Cladosporium fulvum*	avr2	LRR-TM	(12)
	Cf-4	tomato	*Cladosporium fulvum*	avr4	LRR-TM	(13)
	Cf-5	tomato	*Cladosporium fulvum*	avr5	LRR-TM	(14)
	Cf-9	tomato	*Cladosporium fulvum*	avr9	LRR-TM	(15)
	$HS1^{pro-1}$	sugar beet	*Heterodera schactii*	unknown	LRR-TM	(16)
4	Xa21	rice	*Xanthomonas oryzae* pv. oryzae	unknown	LRR-TM-ser/thr protein kinase	(17)
5	Hm1	maize	*Helminthosporium maydice* (race 1)	none	HC toxin reductase	(1)

1997). LRRs can also contain regularly spaced proline, asparagine, or other hydrophobic residues (Bent 1996). Resistance gene proteins may contain multiple LRRs, generally clustered in the N-terminal region. Functionally, LRRs are involved in protein-protein interactions (Jones and Jones 1997; Warren *et al.* 1998) and are therefore prime candidates for the functional ligand binding and recognition domain in the elicitor/receptor model of gene-for-gene resistance (Warren *et al.* 1998). The only LRR protein to have its structure determined by X-ray crystallography is a porcine RNAse inhibitor, which has slightly longer repeats than those found in R genes (Jones and Jones 1997). The functional specificity of the LRRs is thought to reside in the non-conserved residues between the leucine residues, which face outward in the proposed B-helical array, rather than in the leucines themselves which provide the characteristic structure, and are more internally located. Support for the LRR as being important in the disease resistance response is provided by studies of mutant R gene alleles where the LRR has been disrupted, often by only a single amino acid change, resulting in the loss of Avr gene recognition (Baker *et al.* 1997; Bent 1996). LRRs may also participate in the signaling cascade resulting from Avr gene recognition. They have been shown to bind to corresponding ligands in the case of the porcine RNAse inhibitor protein and mammalian hormone receptor proteins (Hammond-Kosack and Jones 1997).

Leucine Zippers are present in a subset of the NBS-LRR R genes, and are also thought to be responsible for protein-protein interactions by the formation of coiled-coil regions (Bent 1996).

Serine/threonine kinase domains have been observed for *Pto* and *Xa21*.

Phosphorylation is known to be a common and important mechanism by which protein activity is controlled (Pawson 1994). Comparison of the *Pto/Prf* genes with *Xa21* indicates that pathogen recognition can be closely associated with phosphorylation, as the kinase *Pto* is required along with the LRR-containing protein *Prf* for resistance, and *Xa21* encodes a transmembrane receptor kinase, with the LRRs extracytoplasmically located, and the kinase domain cytoplasmically located (Song *et al.* 1997).

Nucleotide binding sites (NBS) are often found on LRR-containing R genes. This implies that the binding of nucleotide triphosphate molecules plays a role in protein function. Site specific mutagenesis in these regions has demonstrated the loss of the hypersensitive response in tobacco (*N* gene) and *Arabidopsis* (*RPS2* gene) (Bent 1996).

TIR regions are also found on the NBS-LRR R genes. These are N-terminal domains with similarity to the *Drosophila* Toll protein and the human Interleukin-1 receptor (IL-1R). Speculations on functional similarities of these regions assume that the TIR genes activate defense responses by a mechanism analogous to Toll and IL-1R. In flax, functional analysis of transgenic recombinants of the TIR and LRR domains of 13 *L* locus alleles has shown that allelic differences in specificity can be encoded by either the LRR or TIR domains. Allele pairs with different specificity were found which differed exclusively in either the LRR or TIR domain (Ellis *et al.* 1999).

Some R genes, such as the *Cf* gene of tomato, have transmembrane domains indicating a position bound to a cellular membrane. R gene products anchored to the plasma membrane would be an expected location for a receptor of extracellular pathogens.

1.5.2 R Gene Classes

As more and more R genes were cloned, it was natural to attempt to classify them based on the common motifs, which may reflect their evolutionary origins. Although there is no unanimous classification system among researchers in this field, classification systems generally base classes on the particular motifs present, and the supposed location of the R gene product (Table 1). *Xa21* is always put in a class by itself due to the combination of extracellular LRR and cytosolic Ser/Thr kinase domain, as is *Pto*, being the only non-membrane bound cytosolic kinase. *Hm1* is somewhat of an anomaly in that it is not involved in elicitor recognition or signal transduction, but specifically detoxifies the HC toxin produced by the *Helminthosporium maydis* fungus (Johal and Briggs 1992). The largest class of R genes is the NBS-LRR class, which is often divided into subclasses based on the presence of a LZ or TIR domain. It is interesting to note that one R gene, *RPM1*, confers resistance to *P. syringae* pv. maulicola presenting either the avrRpm1 or avrB elicitor, which shows a slight variation on the strict gene-for-gene model. It is thought that there are two binding sites in the LRR region of *RPM1*, as avrRpm1 and avrB appear to have unrelated structures. Due to many recently cloned R genes falling into existing classes, it may be that the majority of classes of classical R genes have been discovered, although this is by no means certain.

1.6 Comparative Genetic Mapping

Comparative genetic mapping is an invaluable tool for the study of genome organization and evolution (Sadowski and Quiros 1998) and also serves to increase the

efficiency of positional cloning strategies by inferring genetic loci in one species from known loci in another.

Conserved genetic marker order, known as collinearity or synteny, has been observed among groups of plants as high as at the family level, reflecting genomic conservation spanning millions of years. Collinearity has been investigated within Poaceae (Gramineae) (Ahn *et al.* 1993; Ahn and Tanksley 1993; Van Deynze *et al.* 1995), and it was found that marker order from a Triticeae tribe consensus map was conserved along 93, 92 and 94% of the length of rice, maize, and oat linkage maps respectively. Genomic conservation has also been found in the Solanaceae (Bonerbiale *et al.* 1988; Prince *et al.* 1993). Of the 12 linkage groups found in both species, nine were found to have identical marker order along their entire length, with intrachromosomal rearrangements present in the remaining three.

Recent studies have begun to define the extent of collinearity between the model dicot plant *Arabidopsis thaliana* and economically valuable *Brassica* species, thereby determining the feasibility of transferring the abundant genetic and DNA sequence data available from *A. thaliana* to closely related *Brassica* species (Cavell *et al.* 1998). Although the average exon sequence homology between *A. thaliana* and *B. napus* is very high (an average of 87%) (Cavell *et al.* 1998), collinearity of *A. thaliana* markers has been found to span only short stretches (average 8 cM) within the extensively rearranged and duplicated *B. nigra* genome (Lagercrantz 1998), large collinear regions having been disrupted by chromosomal rearrangements (Lagercrantz *et al.* 1996). Internal rearrangement rates have not been reported for *B. napus*, but it has been estimated that

approximately 90 rearrangements have occurred in *B. nigra* since its divergence from *A. thaliana* (Lagercrantz 1998). Due to high chromosomal rearrangement rates in *Brassica* evolution, genetic maps with high marker density are required to obtain a precise representation of intergenomic collinearity (Lagercrantz 1998), and detect chromosomal rearrangements on a small scale (Cavell *et al.* 1998; Lagercrantz *et al.* 1996). Since complete *Brassica* genome maps still have relatively large average distances between markers (Sadowski *et al.* 1996), much *A. thaliana/Brassica* comparative mapping has focused on microsynteny; mapping markers from short, well characterized and marker-dense segments of *A. thaliana* chromosomes.

Analysis of *A. thaliana/B. napus* collinearity along a 30 cM, 7.5-Mbp segment of *A. thaliana* chromosome four with 10 marker loci revealed four perfectly collinear regions, and an additional two with identical internal inversions in three *B. napus* homoeologous chromosome pairs (Cavell *et al.* 1998), reflecting the hypothesized hexaploid ancestry of the *B.napus* genome. Investigation of a cluster of five physically mapped genes on *A. thaliana* chromosome three revealed one completely collinear region in each of the three *Brassica* genomes (A, B, and C genomes from *B. rapa*, *B. nigra* and *B. oleracea* respectively), as well as partial clusters on one (potentially two for *B. oleracea*) other linkage group (Sadowski *et al.* 1996). Conservation of marker order in *Brassica* could not be assessed due to lack of recombination within the gene cluster (Sadowski *et al.* 1996). A cluster of six genes on *A. thaliana* chromosome 4, including the disease resistance gene RPS2, was found intact on *B. nigra* linkage group B1, and incomplete on linkage groups B4 , B2, B8, and B3 (Sadowski and Quiros 1998). It generally appears that chromosomes

ancestral to *Brassica* and *Arabidopsis* underwent both duplication and re-patterning in the *Brassica* lineage, leading to higher loci copy number, larger genomes, and scattered homoeologous locus distribution.

Collinearity between *Brassica* and *A. thaliana* allows candidate loci to be identified in *Brassica* based on homology to characterized loci in *A. thaliana*. Two loci influencing flowering time in *B. nigra* were identified by homology to the *A. thaliana* flowering time gene *CO* (Lagercrantz *et al.* 1996), illustrating the power of comparative mapping by utilizing *A. thaliana* gene markers to identify homoeologues in collinear regions of related genomes.

This candidate gene approach has also been used to identify loci with potential disease resistance activity based on sequence homology to conserved protein motifs in previously cloned plant disease resistance genes. Cloned disease resistance genes from wide ranging taxa and against various pathogenic organisms (i.e. bacteria, fungi, viruses, nematodes) share structural or amino acid motifs. Homology among disease resistance gene products implies conservation of defense response mechanisms (Baker *et al.* 1997; Bent 1996), and presents a means to isolate DNA fragments with putative disease resistance function (Kanazin *et al.* 1996).

1.7 Exploitation of R Gene Homology in Conserved Motifs

1.7.1 PCR Amplification of R Gene Homologues Using Degenerate Primers

Although low overall DNA sequence homology among R genes precludes using them directly as probes to identify R gene homologues by Southern hybridization to genomic

DNA (Leister *et al.* 1996; Yu *et al.* 1996), PCR with primers targeted to the most conserved LRR or NBS domains has met with success in amplifying R-gene homologues (Kanazin *et al.* 1996; Leister *et al.* 1996; Ohmori *et al.* 1998; Yu *et al.* 1996). Yu et al. (1996), using degenerate PCR primers for the NBS domains of the tobacco R-gene *N* (conferring resistance to tobacco mosaic virus) and the *A. thaliana* R-gene *RPS2* (conferring resistance to the bacterial pathogen *Pseudomonas syringae* pv maulicola), amplified fragments in 11 classes (multigene families) with high homology to *N*, *L6*, or *RPS2* along their entire length. The NBS clone classes showed 19, 25.5, and 29.5% average similarity to the NBS regions of *RPS2*, *N*, and *L6* respectively, which is comparable to the sequence similarity observed between *RPS2*, *N*, and *L6* themselves (Yu *et al.* 1996). Using RFLP mapping with near isogenic lines (NILs) of soybean, five of the 11 classes coincided with known disease resistance gene clusters, although no individual NBS clones were identified as segregating with known disease resistance genes (Yu *et al.* 1996). A similar experiment done by Kanazin et al. (1996) identified resistance gene-like sequences which they called R gene analogs (RGAs). Again, RGAs were grouped into classes, and after mapping a member of each class, RGAs were found near known resistance loci. The RGAs were also found to have a clustered genomic organization, which may be a significant reflection of the clustering found in many well studied R gene loci (Kanazin *et al.* 1996). This clustering of R genes suggests common genetic mechanisms in their evolution.

PCR primers based on the NBS sequence of *N* and *RPS2* were also used to amplify and clone fragments from potato (Leister *et al.* 1996). In this case, R gene homologue markers

were found to cosegregate with the *Gro1* nematode resistance gene and the *R7 P. infestans* gene in 100 and 96 individuals respectively (Leister *et al.* 1996). Similarly amplified products in tomato were homologous to LRR + NBS R genes (Ohmori *et al.* 1998). Two of these tomato clones (TC10.1 and TC11.2) were found to be 98.7% and 93.9% homologous to potato DNA fragments (St124 and St332 respectively) amplified by Leister et al. (1996) (Ohmori *et al.* 1998), reflecting the genomic conservation between these two species (Bonerbiale *et al.* 1988).

This candidate gene approach to discovering new R genes is not without problems. Degenerate PCR primers may produce products of no functional significance (Leister *et al.* 1996), and there are other NBS and LRR- containing genes not involved in disease resistance, although as more is learned about specific R gene signatures, this should become less of a problem. Silenced pseudogenes in multigene resistance families may be a prime source of polymorphism in RFLP mapping due to their freedom to accumulate mutations. Due to this, R gene homologues may be most useful in landing markers close to disease resistance genes rather than identifying functional R genes directly (Kanazin *et al.* 1996). There is also the problem associated with methods to distinguish specificities encoded by different members of tightly linked multigene families (Yu *et al.* 1996).

The use of multiple common motifs may increase the likelihood of isolating homologues with disease resistance function. The largest class of R genes contains both NBS and LRR regions, which was exploited with PCR primers designed for each region to amplify fragments with homology to both conserved motifs (Joyeux *et al.* 1999). R gene homologues were amplified from *B. napus* cv. Quantum, and many have been

mapped in this research using the N-fo-61-9 *B. napus* population.

1.7.2 Utilizing Expressed Sequence Tags in the R Gene Candidate Search

Utilization of expressed sequence tags (ESTs) is a fast way to obtain information on the coding regions of an organism's genome (Newman *et al.* 1994). ESTs are derived from single-pass, random sequencing of cDNAs from various libraries. Generally, anywhere from 250 to 400 base pairs are sequenced for each EST, and are entered into the EST database (dbEST) (Boguski *et al.* 1993) which contains over 3,000,000 sequences from over 180 organisms (www.ncbi.nih.gov/dbEST/ dbEST_summary.html), with almost 46,000 from *A. thaliana* alone.

Due to common functionality of conserved amino acid sequence domains, one can often assign putative function to an EST based on amino acid or DNA sequence homology to a known gene, although correct identification of function relies on extensive genetic and biochemical investigation (Cooke *et al.* 1996; Leister *et al.* 1996; Newman *et al.* 1994). ESTs can be identified by homology to a characterized gene sequence, resulting in a population of tags for expressed genes which have structural similarity to the gene originally used as the search template. This approach was taken by Botella et al. (1997) to search for ESTs which show homology to any of *Cf-9*, *Xa21*, *RPM1*, *RPS2*, *L6*, *N*, *RPP5* or *Pto*. In order to have significant sequence similarity, only ESTs with BLAST scores >80 were considered (Newman *et al.* 1994). Ninety-four *A. thaliana* EST sequences representing 62 non-redundant clones were identified from the R-gene homology search, and 42 were mapped in *A. thaliana* by assignment to physically mapped yeast artificial

chromosomes (YACs) using PCR, or by molecular marker segregation in *A. thaliana* recombinant inbred (RI) lines. Due to the size of the YACs, the map positions are generally imprecise, representing a 3 cM interval on average, and may be incorrect due to the presence of multigene families with highly related members and/or chimeric YACs (Botella *et al.* 1997). The database search identified ESTs that were highly similar to *RPS2*, *RPM1*, and *RPP5*, indicating that the possibility of identifying unknown R genes exists using this approach. R genes are known to be expressed at low levels in uninfected tissues, and should be well represented in the dbEST (Botella *et al.* 1997; Hammond-Kosack and Jones 1997). Even though this is the case, cDNA libraries from infected tissues may be enriched in R gene, as well as other defense related mRNAs. As many elements of the disease response signal cascade are common for different pathogens, there would be no need to construct libraries from every plant/pathogen interaction to identify these signal transduction intermediates. *A. thaliana* ESTs have been sequenced with homology to genes not yet described in higher plants (Newman *et al.* 1994). This provides the opportunity to isolate these genes by using the ESTs as probes to genomic libraries. It has also been proposed that ESTs could somehow be used in directed mutation of their corresponding genes. This directed mutagenesis, if developed, would be a useful tool in generating mutants, as there would be an instant marker for the gene, and some limited sequence information (Newman *et al.* 1994).

Due to the close evolutionary relationship between *Arabidopsis* and *Brassica* (both in the Brassicaceae), ESTs from *Arabidopsis* will hybridize to homologous genomic regions of *Brassica*, and the positions of the same marker in both genera can be compared.

1.8 Evolution of R Gene Loci

As R genes evolve to recognize pathogen avirulence genes, pathogens evolve to evade detection. Pathogen evolution from virulent to avirulent requires a loss of function mutation as explained by the gene-for-gene theory. Avr gene mutations occur at high frequencies in nature (McDowell *et al.* 1998), and combined with the fact that pathogen populations sizes are much larger, and have a faster reproduction rate, it is difficult to envision how R gene loci can remain in the arms race, as a gain of function mutation would be required to produce an R gene with a new specificity (Bennetzen and Hulbert 1992; Staskawicz *et al.* 1995). There must be other mechanisms at work which allow plants to generate novel R genes to maintain resistance against constantly changing pathogens.

The original ancestors of R genes were most likely endogenous recognition or signaling proteins required for normal plant growth and development. For example, *A. thaliana* genes *ERECTA* and *CLAVATA* are floral organ morphology genes which have homology to the extracellular LRR region of *Xa21*, and are known to bind ligands in intercellular communication.

The organization of R genes within their respective genomes gives evidence as to their evolution. Two common types of organization observed for R gene loci are non-allelic R gene families with different race specificities, often in arrays of tandem repeats, or a single gene with multiple alleles at the same locus. These are exemplified by the *M* and *L* loci, which have the two respective types of organization, yet have 70-90% nucleotide sequence identity (Staskawicz *et al.* 1995). In the case of *RPM1*, there is a single gene

present in resistant lines and absent in susceptible lines. R genes can also be loosely clustered, as in the major recognition complex (MRC) regions of *Arabidopsis* (Hammond-Kosack and Jones 1997).

In complex R gene loci (loci with linked tandem repeats), the linked genes may show a relationship where they encode multiple disease resistance functions, as in the linked *Pto*, *Fen*, and *Prf* cluster of tomato, or multiple disease resistance specificities of similar function as in the complex *N*, *M*, *RPS2*, *Cf* and *Rp1* loci (Bennetzen and Hulbert 1992; Dangl 1995). Functional R genes have also been found in clusters of homologues which appear to be truncated or incomplete sequences not believed to be transcribed. It has been proposed that the tandem linkage of R genes and R gene homologues is indicative of their method of evolution.

Unequal crossing over events at these loci are thought to cause an increase in diversity, and additional R genes to be formed at a complex R locus (Bennetzen and Hulbert 1992; Jones and Jones 1997). Generation of altered specificity related to unequal crossing over has been seen in the loss of rust resistance associated with unequal recombination at the *Rp1* locus of maize (Bennetzen and Hulbert 1992). The structure of the susceptible *RPP8*-Col locus was generated by unequal crossing over of linked, non-allelic genes, but is believed to recognize an unknown pathogen, as the chimeric protein is expressed (McDowell *et al.* 1998). The amino acid substitution pattern of the *RPP8* gene and one homologue in Columbia and Landsberg ecotypes indicate positive diversifying selection is acting on these loci. The number of non-synonymous substitutions (Ka) was greater than the number of synonymous substitutions (Ks), leading to a Ka/Ks ration of greater

than one, indicating non-conservative evolution. The LRR regions of the three sequences showed the greatest amino acid divergence, supporting the idea of LRRs harboring Avr gene elicitor specificity (McDowell *et al.* 1998).

Adaptive selection has also resulted in divergence of the rice *Xa21* and *Xa21D* genes. *Xa21D* is one of a number of linked homologues in the *Xa21* gene family, and encodes a truncated protein similar to a protein kinase without the kinase domain. It has no transmembrane domain, and based on similarity to the *S* locus glycoprotein and polygalacturonase inhibiting protein of *Brassica*, is thought to be secreted into the extracellular matrix and bind to the Avr encoded ligand. *Xa21D* is the only linked homologue to show a resistance phenotype, conferring only partial resistance to *X. oryzae* (Wang *et al.* 1998).

The molecular evolution of the *RPS2* gene of *A. thaliana* has been closely examined by Caicedo et al. (1999) in order to evaluate levels of polymorphism at this locus, evolutionary relationships among alleles, possible selection pressure, and the relationship between sequence and disease resistance response. Among 17 *A. thaliana* accessions, 36 polymorphic sites were found, with a disproportionate amount located in the LRR region, and twice the number of non-conservative amino acid substitutions as conservative. In fact, the *RPS2* gene was determined to have one of the highest occurrences of sequence polymorphism found among plant genes to date (Caicedo *et al.* 1999), which would be expected in a gene required to keep up with rapid pathogen evolution. The resistant haplotypes of *RPS2* were all closely related, with only a few amino acid changes between them, but susceptible haplotypes exhibited widely divergent sequences. R genes can only

tolerate a certain level of mutation before losing function and reverting to a susceptible form (Caicedo *et al.* 1999).

In the tomato *Cf-5* cluster, six of the seven linked homologues show variation in the numbers of LRRs in the N-terminal region, while the C-terminal portions are highly similar. Also, a recombination between *Cf-2* and *Cf-5* loci was shown to be intragenic. (Dixon *et al.* 1998). Intragenic unequal recombination could expand or contract the number of LRRs, generating diversity and altered specificity (Fletcher 1994).

Xa21 belongs to a multigene family with at least eight members, which map to a locus on rice chromosome 11, linked to nine known disease resistance specificities. Transposition, duplication, and recombination have all contributed to the diversity at this complex locus. Fifteen sequences designated as transposable elements based on the terminal inverted repeats were found in members of the *Xa21* family (Song *et al.* 1997). Insertion and excision of transposons would generate polymorphism at complex loci, and would be another mechanism for generation of new R gene alleles. A highly conserved (HC) GC-rich region was present in all family members, and may mediate intragenic recombination events, as several recombination breakpoints were clustered in this region (Song *et al.* 1997). This HC region has not been described for any other R-genes. An ancient duplication event and subsequent diversification can be seen in the classification of the *Xa21* gene family members into two distinct groups.

A transposon has recently been implicated in the evolution of the *A. thaliana RPS5* gene. *RFL1* is a gene with unknown function linked to *RPS5*, and with 74% nucleotide identity in the Col-0 accession, and is also present in the Ler-0 accession which lacks

RPS5. Sequencing of this region in both accessions lead to the conclusion that *RPS5* was lost in Ler-0 not from a recombination between the two, but from an instance of non-homologous end-joining, perhaps triggered from insertion or deletion of the Tag2 transposon. *RPS5* was not found to be simply deleted in Ler-0, but replaced with 3 kb of sequence with highest similarity to the Ac transposable element of maize (Henk *et al.* 1999). *RPS5* was likely deleted during repair of a double-strand break, which may have been caused by Tag2. In non-homologous end-joining, a double stranded DNA break is followed by exonuclease digestion of the two ends. The ends are subsequently re-joined by an unknown mechanism, often with 'filler' DNA copied from regions a few kb away. Non-homologous end joining is the main repair mechanism in plants for double stranded breaks, and could be an important polymorphism generator at R gene loci (Henk *et al.* 1999).

With all of these mechanisms in place to generate novel R genes, is there any limit, or advantage in limiting the number of R genes in any single plant? The fact that R genes have been deleted numerous independent times (Grant *et al.* 1998; Henk *et al.* 1999; McDowell *et al.* 1998) indicates that there could be a negative fitness cost associated with R genes in the absence of selection pressure. Both the resistant and null susceptible *RPM1* genotypes have proliferated in geographically widespread populations, indicating that variation in pathogen pressure is maintaining both alleles (Grant *et al.* 1998). A proposed mechanism for R gene cost is the low level, inappropriate induction of the disease response, resulting in metabolic shifting of resources to the production of unnecessary compounds and/or cell death in the HR (Grant *et al.* 1998; Henk *et al.* 1999).

Rapid rearrangement of NBS-LRR homologues has been observed across monocot genomes (Leister *et al.* 1998). The high level of genomic synteny among the cereals is maintained by the detrimental effects of ectopic recombination. A different mechanism has been implied to exist for cereal R gene evolution, as their organization does not reflect this strict interspecific conservation. Comparative mapping of rice and barley R gene homologues in rice, barley, and foxtail millet revealed interspecific variation in copy number, clusters of highly dissimilar R gene homologues, frequent non-syntenic map locations, and a lack of interspecific hybridization signals for some of the homologues. Lack of cross hybridization could be the result of rapid sequence divergence, or specific and unique ectopic recombination events at R gene loci in each of the three monocot species (Leister *et al.* 1998). All these lines of evidence support a higher evolution rate of R genes than is evident in the rest of the cereal genome.

1.9 *Brassica* Genome Relationships

The genus *Brassica* includes a wide variety of crop plants including cabbage, cauliflower, broccoli, turnip and mustard. Oilseed *Brassica* species (*B. napus* and *B. rapa*) provide more than 12% of the world's edible oil, and Canada is the fourth largest producing country. The genomic relationships among the economically important Brassicas is well known. The three diploid species (*B. rapa*, *B. nigra*, *B. oleracea*) have hybridized in all combinations to produce the three amphidiploid species (*B. carinata*, *B. juncea* and *B. napus*) (Figure 3).

Based on comparative RFLP mapping, it was observed that the three diploid species have collinear regions covering nearly the whole of each genome, but with extensive, large chromosomal rearrangements (Lagercrantz and Lydiate 1996). The ancestral genome

has been proposed to be n=6, and all three diploid species have inherited rearranged, but

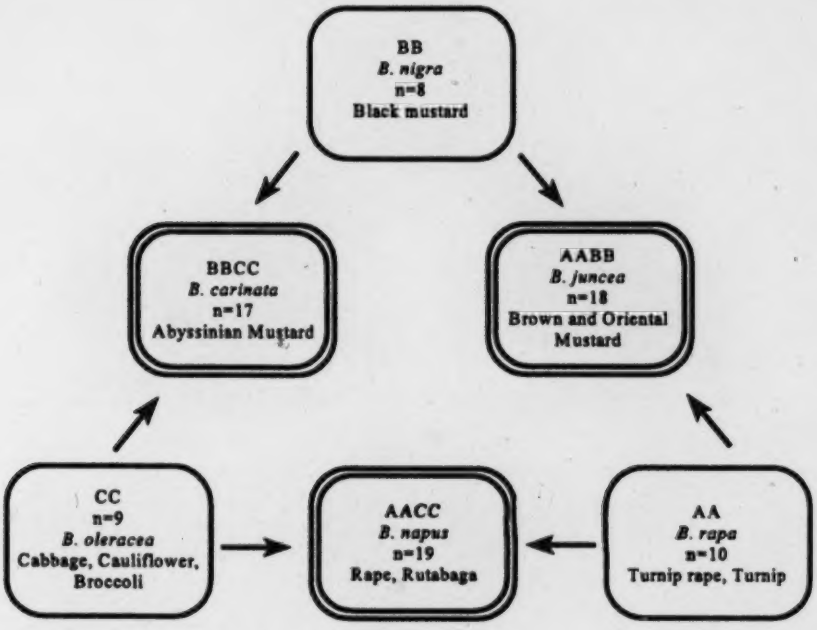


Figure 3. Genomic relationships between the diploid and amphidiploid *Brassica* species. After U (1935).

essentially complete copies of this genome. In general, the collinear genomic segments appear to be triplicated in the diploid species, most likely from a hexaploid ancestral genome (Lagercrantz and Lydiate 1996).

The genome of *A. thaliana* is generally free from large duplicated segments (Chang *et*

al. 1988; Nam *et al.* 1989) although duplicated loci are known to be present (McGrath *et al.* 1993). The genome is similar in size to the triplicated unit genome of *B. nigra*, and has been proposed to be similar to the hypothesized hexaploid ancestor of the Brassicas (Lagercrantz and Lydiate 1996). If this is the case, then the large number of rearrangements from the ancestral genome has taken place in the last 10 million years, a much more rapid rate of evolution than is seen in the evolution of the cereal genomes (Lagercrantz 1998).

The genomes of the diploid ancestors of *B. napus* (*B. oleracea* and *B. rapa*) have remained essentially unchanged since the interspecific hybridization event, evidenced by the normal meiotic pairing of chromosomes in a cross between a natural and re-synthesized *B. napus* (Parkin and Lydiate 1997; Parkin *et al.* 1995).

In a study to investigate the genomic relationship of *A. thaliana* and *B. nigra*, it was found that each *B. nigra* linkage group (except one) hybridized with probes from all 5 *A. thaliana* linkage groups, further demonstrating the high numbers of recombination events in the *Brassica* lineage (Lagercrantz 1998). The average length of conserved segments between *A. thaliana* and *B. nigra* was found to be 8 cM. This allows genes mapped to a <10 cM interval within a *Brassica* genome to potentially have a collinear location determined in *A. thaliana*. This allows exploitation of the high density of markers on the *A. thaliana* map, as well as *A. thaliana* YACs, BACs, and ESTs (Lagercrantz 1998).

1.10 Mapped *Brassica* R Genes

Compared to species such as wheat, with over 100 mapped disease resistance loci,

resistance phenotype mapping in the Brassicas is in its infancy. No R genes have yet been cloned from any *Brassica* species. Most of the disease resistance mapping has been done with the oilseed Brassicas, trying to map sources of Blackleg (*Leptosphaeria maculans*) resistance. The major gene of an adult blackleg resistance quantitative trait locus (QTL) from the french *B. napus* cultivar Cresor (*LmFr*$_1$) has been mapped to a 10 cM interval by Dion et al. (1995). Another *L. maculans* resistance locus, from the cultivar Major, has been mapped to linkage group 6 in the Major x Stellar F_1 DH mapping population (Ferreira *et al.* 1995a) with minor genes on linkage group 17. It has been designated *LEM1*. Mayerhofer et al. (1997) mapped a third source of *L. maculans* resistance from the Australian cultivars Shiralee and Maluka, designated *LmR1*. *LEM1* and *LmR1* were found to map to different positions on the same linkage group, indicating possible clustering of R genes on this linkage group. A white rust (*Albugo candida*) resistance locus (*ACA*) has been mapped in *B. napus* by Ferreira et al. (1995b) located on linkage group 9 of the Major x Stellar mapping population. Mapping resistance genes is the first step in map-based cloning strategies, and undoubtedly one or more of these genes will be cloned in the near future.

1.11 Objectives of This Research

The objectives of this research were to identify and map disease resistance candidate genes in both canola (*Brassica napus*) and *Arabidopsis thaliana*. This will facilitate present and future map-based R gene cloning efforts by potentially allowing the omission of time consuming fine mapping around disease resistance loci and probing of genomic

YAC or BAC libraries. A resistance phenotype mapped very close to a candidate resistance gene will have a potentially functional marker already present. The DNA probe used to generate the marker can then be examined for disease resistance function by complementation. The addition of candidate R genes to the *B. napus* map will be particularly important, as it is an economically valuable crop species.

To locate candidate R genes, 50 R-ESTs identified by Botella et al. (1997) were obtained and used as RFLP markers in both *A. thaliana* and *Brassica napus*. Twenty-two cloned *Brassica* R-gene homologues provided by Alex Joyeux, a *Pto* homologue provided by Bret Kennedy and an $HS1^{Pro-1}$ homologue provided by Dr. Dwayne Hegedus were also mapped in *B. napus* to provide additional loci (Joyeux *et al.* 1999).

Another goal of this research was to further the identification of collinear regions between the genomes of *B. napus* and *A. thaliana*. This allows the large amount of sequence data available for *A. thaliana* to be used by *Brassica* researchers.

II. Materials Methods

2.1 Plant Mapping Populations

The *B. napus* mapping population consisted of 30 doubled haploid (DH) lines of the highly polymorphic N-fo-61-9 population (Figure 4). Parents of this population are a DH British winter cultivar (N-o-9) derived from microspore culture (P. Capitain and R. Jennaway, Cambridge Plant Breeders, Thriplow SG8 7RE, U.K.) and a resynthesized

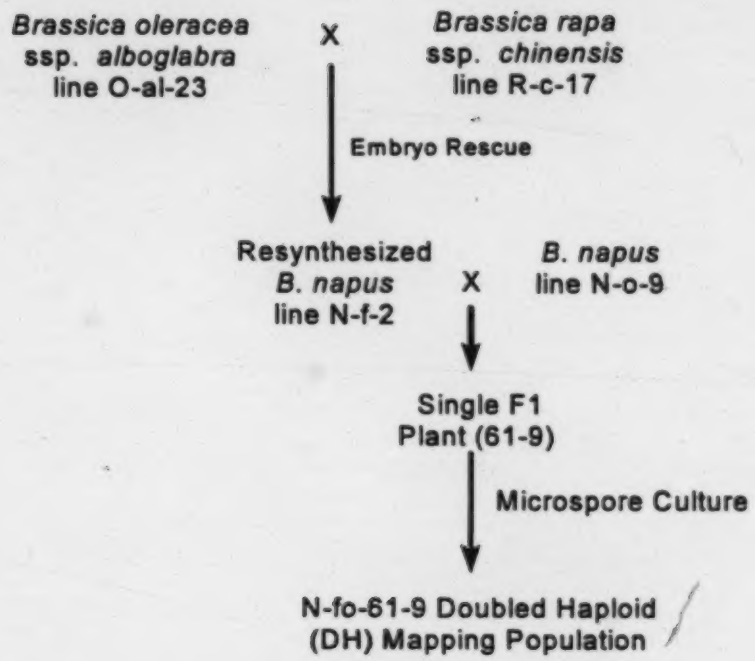


Figure 4. Generation of the N-fo-61-9 Brassica napus mapping population used for mapping R-ESTs and Brassica R gene homologues. Parkin et al. (1995).

B. napus (N-f-2) produced from the cross *B. rapa* ssp. chinensis (R-c-17) x *B. oleracea* ssp. alboglabra (O-al-26) followed by embryo rescue (T. Hodgkin, Scottish Crop Research Institute, Dundee, DD2 5DA, U.K.) (Parkin *et al.* 1995). The complete N-fo-61-9 population consists of 50 DH *B. napus* lines produced via microspore culture (Chuong and Beversdorf 1985) from a single progeny plant of the N-o-9 x N-f-2 cross. This population was kindly provided with support from Zeneca Seeds.

The *Arabidopsis thaliana* mapping population consisted of 30 recombinant inbred (RI) lines derived from a Columbia x Landsberg *erecta* ecotype cross (Lister and Dean 1993). Thirty lines is adequate for doing mapping at the scale of the entire genome if individuals are selected with recombinations at set intervals. Many individuals in a mapping population will have redundant recombination information and by selecting individuals with the most informative recombinations, marker order can be assigned with fewer individuals. For example, Joyeux et al. (1999) were able to arrive at the same marker order for *B. napus* linkage group 1 (cv. Major x Stellar cross) using a subset of 23 informative individuals from the original mapping population of 105.

2.2 Genomic DNA Extraction and Southern Hybridization

Genomic DNA was extracted from 0.5 g of freeze dried leaf tissue as described by Sharpe et al. (1995), using "Kirby mix" as the initial extraction buffer, followed by phenol/chloroform extraction, isopropanol precipitation, RNAse treatment, a second phenol/chloroform extraction and a final isopropanol precipitation. Restriction endonuclease digestions were done in 4 mM Spermidine, 1x appropriate SuRE/Cut Buffer (Boeringer Mannheim), 10 μg/μl Bovine Serum Albumin (BSA) (*A. thaliana* digests only), 15 μg genomic *B. napus* or 4 μg genomic *A. thaliana* DNA, and 2 units enzyme/μg DNA; digested overnight at 37°C. *Brassica napus* genomic DNA was digested with either EcoRI or EcoRV; *A. thaliana* DNA with ClaI, CfoI, DraI, EcoRI, EcoRV, HpaII or ScaI. The highly polymorphic nature of the N-fo-61-9 *B. napus* population allowed numerous RFLPs to be detected between the parents with only 2

restriction enzymes.

Gel electrophoresis of the digested DNA was done with 0.8% agarose gels in 1x TAE buffer overnight at 1.5 V/cm, until the dye front was 12 cm from the wells. Alkaline transfer and Southern hybridization were done according to Sharpe et al. (1995). Filters were washed twice in 2X SSC, 0.1% SDS at 65°C, and once at room temperature, wrapped in cling film, and placed against film (Kodak X-OMAT AR) for autoradiography at -80°C.

2.3 RFLP Probes

R-EST probes used for RFLP mapping were a subset of expressed sequence tags (ESTs) screened from the *Arabidopsis thaliana* database (http://genome-www.stanford.edu/ *Arabidopsis*/) selected based on nucleotide sequence homology to one of 8 cloned plant disease resistance genes (see introduction) (Botella *et al.* 1997). A BLAST score of >80 was used as the criterion for sequence similarity significance (Newman *et al.* 1994). R-EST probes were obtained from the *Arabidopsis* Biological Resource Centre in Ohio, and from Dr. Patrick Dao and Gabriel Philipps of the Laboratoire de Gênêtique Moleculaire des Plantes, Grenoble, France. They were provided as inserts in pBluescript SK-, pZL1, lambda Zip-Lox, pSHlox-1, or Lambda Zap II vectors in *E. coli*. Single colonies were isolated from LB+Amp. agar plates and cultured in LB+Amp. medium overnight in a shaking incubator at 37°C. A glycerol stock was prepared from each culture (850 μl culture to 150 μl glycerol) and stored at -80°C. *Brassica* R gene homologue probes were previously cloned from *B. napus* cv. Quantum

using primers homologous to LRR and NBS motifs of the *A. thaliana* RPS2 gene, and were provided by A. Joyeux (Joyeux et al. 1999). Bacterial cultures of *Brassica* homologues were inoculated from glycerol stocks. *B. napus Pto* (H-3) and *HS1^{pro-1}* (HS1) homologous probes were provided by B. Kennedy and D. Hegedus respectively.

Vector plasmids were isolated with a standard mini-prep procedure for use as template in PCR reactions to amplify the inserts for use as probes in Southern hybridization.

2.4 PCR Conditions

Cloned *Brassica* R-gene homologues and R-EST probe inserts were amplified in 25 μl polymerase chain reactions containing 50 mM KCl, 10 mM Tris-HCl (pH 8.0), 0.01% gelatin, 2.5 mM $MgCl_2$, 0.2 mM each of dATP, dCTP, dTTP, and dGTP, 0.4μM primer (T3, T7, or SP6), 0.1 U Taq polymerase, and 25 ng of R-EST template DNA. Reaction parameters were 94°C for 2 min; 35 cycles of 94°C for 1 min, 65°C for 1 min and 72°C for 1 min; 72°C for 5 min, in a Gene Amp 9600 thermocycler (Perkin-Elmer). PCR products were passed through Sepharose CL-6B (Pharmacia) spin columns to remove unincorporated dNTPs before visualization and ^{32}P labeling for use in Southern hybridization. PCR product was radioactively labeled with a random priming method (N6 primers) and Klenow fragment (Pharmacia) extension in the presence of ^{32}P-dCTP.

2.5 Linkage Analysis

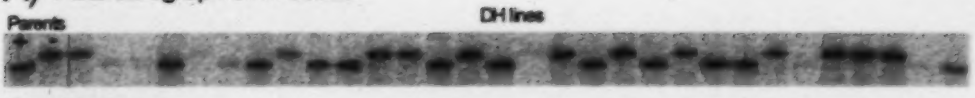
Assignment of RFLP loci to the 19 established *B. napus* linkage groups was done both manually and with MAPMAKER/EXP version 3.0 (Lander *et al.* 1987). Data were entered

into MAPMAKER after specifying that +=A, - =B, and 0= - at the beginning of the data file. Markers previously mapped with the N-fo-61-9 population were anchored to specific linkage groups with MAPMAKER, and each new marker was located within this well established framework (Cavell *et al.* 1998; Parkin and Lydiate 1997; Parkin *et al.* 1995). Minimum LOD (logarithm of odds) scores of 3.0 were used to initially assign loci to linkage groups, and three-point and multipoint analyses were used to determine the most likely map position within the linkage group. Manual fine adjustment was done to minimize double crossovers within short map intervals. Double crossovers should be rare over short chromosomal distances due to individuals of the N-fo-61-9 population having been derived from a single round of meiosis in the source F1 plant (Sharpe *et al.* 1995).

2.6 Brief Overview of RFLP Map Generation

The first step towards the generation of a genetic map from RFLP markers is to look at the autoradiograph of the labeled probe hybridized to the digested genomic DNA of the two parents and each individual member of the mapping population. Polymorphic alleles at a locus between parents will generally segregate among the progeny in a disomic pattern; each individual inheriting one allele or the other in the case of doubled haploids and recombinant inbred lines. As data from more and more probes are collected, it will be observed that certain pairs of markers are inherited in a similar fashion among members of the mapping population (Figure 5). This is far more likely to occur if the markers are linked than if the markers were inherited randomly. The probability of markers being linked is given by the logarithm of odds score (LOD) which is the ratio of the probability

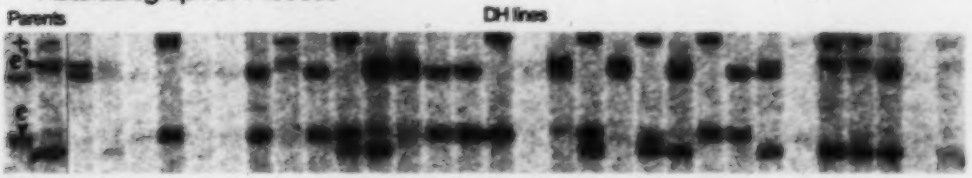
A) Autoradiograph of T75662a



Scoring Data*

T75662a -++0-++-++--+-+--+-+-++-+---++

Autoradiograph of T43968e*



Scoring Data*

T43968e --+0-++-++--+++--+-+-++-+---+-

B) Integrated Scoring Data

RFLP Loci																																	
*pW137b	-	-	+	-	-	+	-	-	+	-	-	-	+	+	+	-	-	+	-	+	-	+	+	-	+	-	-	-	-	+	-
*T43968e	-	-	+	0	-	+	+	-	+	+	-	-	+	+	+	-	-	+	-	+	-	+	+	-	+	-	-	-	-	+	-
*pN66a	-	-	+	-	-	0	+	0	+	+	-	-	+	-	+	0	-	0	-	+	-	+	+	-	+	-	-	-	-	+	-
*APSb	-	-	+	+	-	0	+	0	+	+	-	-	+	-	+	0	-	0	-	+	-	+	+	-	+	-	-	-	-	+	+
*ACYLa	-	+	+	+	-	+	+	-	+	+	-	-	+	-	+	-	-	+	-	+	-	+	+	-	+	-	-	-	-	+	+
*pN53dNM	-	+	+	+	-	+	+	-	+	+	-	-	+	-	+	-	-	+	-	+	-	+	+	-	+	-	-	-	-	+	+
*T75662a	-	+	+	0	-	+	+	-	+	+	-	-	+	-	+	-	-	+	-	+	-	+	+	-	+	-	-	-	-	+	+

C) Mapmaker Output

```
Markers                Distance
281  pW137b              6.7 cM
282  T43968e             3.4 cM
283  pN66a               6.8 cM
284  APSb                3.4 cM
285  ACYLa               0.0 cM
293  pN53dNM             0.0 cM
294  T75662a           ----------
                        20.4 cM   7 markers   log-likelihood= -38.44
```

Figure 5. A) Autoradiographs for R-EST probes T75662 and T43968 on the N-fo-61-9 *B. napus* DH mapping population. + allele inherited from N-o-9 parent; - allele inherited from resynthesized N-f-2 parent; 0 missing data point. B) Inheritance pattern of the parental alleles at the T75662a and T43968 loci is consistent with that of the flanking markers. Note the chromosomal recombination events between the two R-EST marker loci in the individual DH lines marked with arrows. C) Linkage map of this small region of N5 generated with MAPMAKER. This particular order of markers is $10^{38.4}$ times more likely than any other order. * The fourth lane from the left in the DH mapping population is scored as '0' because this individual was from a different population in this particular set of filters.

that a marker is linked to the probability that it is not linked. A marker with a LOD score of 3.0 (10^3 times more likely to be linked to its neighbors than unlinked) is generally considered to be strong evidence for linkage.

2.7 Atypical RFLP Patterns

As mentioned previously, most marker loci segregated in a disomic fashion, with the DH or RI plants inheriting only one allele of a pair present in the parents (ie. none inheriting both, or neither allele). This is consistent with normal bivalent formation in the F1 plant, with the 10 *B. rapa* chromosomes exclusively pairing with the 10 A genome chromosomes of the winter *B. napus* parent (N-o-9), and likewise for the 9 chromosomes of the C genome of *B. oleracea* (Sharpe *et al.* 1995) (Figure 6).

In the N-fo-61-9 population, not all markers segregated in a disomic pattern. In certain regions of the genome, one allele from the resynthesized parent (N-f-2) was absent from individuals of the mapping population, and a second (homoeologous) allele was present in a high number of the mapping population individuals and in a varying dose of 0,1, or 2 copies. This pattern is indicative of duplication of chromosomal segments with concurrent replacement of homoeologous segments (Parkin *et al.* 1995). This is most likely due to homoeologous recombination in the N-f-2 parent, leading to non-reciprocal homoeologous translocations in the gamete that N-f-2 donated to the N61-9 F1 plant from which the DH mapping population was made (Parkin *et al.* 1995).

Tetrasomic inheritance can be seen in chromosomes N2 and N12 of *B. napus*. These chromosomes are made up of almost equal sections of the A and C genomes, so each of

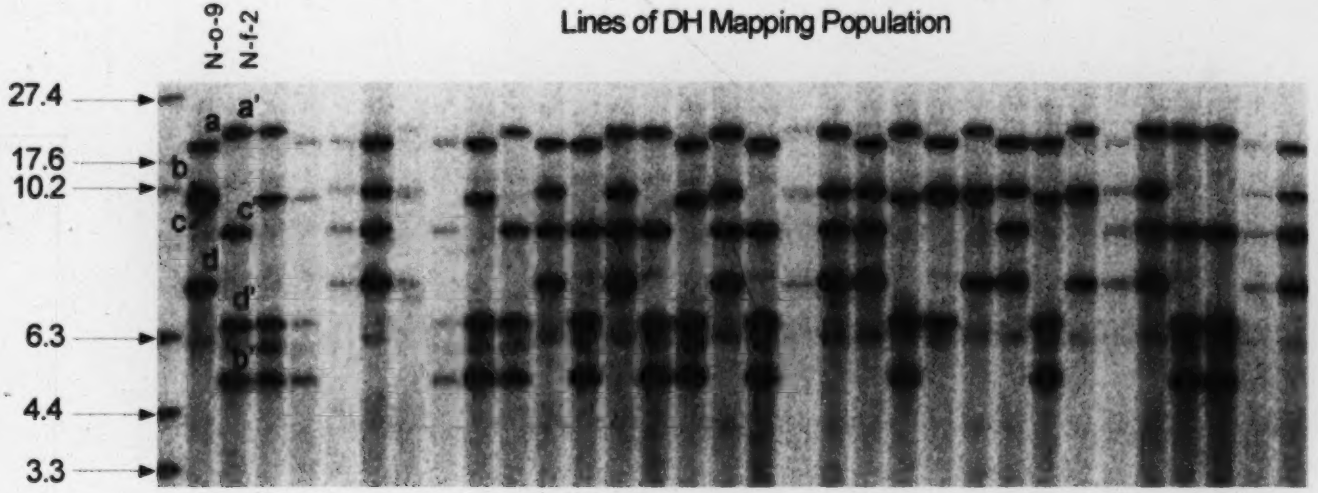


Figure 6. RFLP patterns typical of disomic inheritance. Each individual of the N-fo-61-9 DH mapping population received a single parental allele at each polymorphic locus, indicating normal homologous chromosome pairing at meiosis. Each locus is designated with a lower case letter, and alleles from the resynthesized parent are designated with (’). Numbers to the left are size standards in kilobase pairs (kb).

N2 or N12 from the N-o-9 parent could pair with either N2 or N12 from the N-f-2 parent. This results in each DH line of the population having received any 2 of the 4 alleles from the 2 homoeologous loci (Parkin *et al.* 1995).

III. Results and Discussion

3.1 Genetic Mapping of *Arabidopsis* R-ESTs

3.1.1 R-EST Loci Detected in A. thaliana and B. napus

Of 50 *Arabidopsis* R-ESTs chosen as RFLP probes: five (H37061, N65692, N96711, R29891 and Z34772) hybridized to *B. napus* genomic DNA but detected no polymorphic bands between the parental lines, N-o-9 and N-f-2; one (R64749) proved to be an incorrect clone (leaving a total of 49 R-ESTs); and four (F20108, T04109, T21150 and T20493) proved to be redundant since they detected RFLP patterns identical to probes previously used in the study (Table 2). Botella et al. (1997) had previously indicated that F20107 and F20108 were from the same clone and had placed T21150 in a contig with R89998; however T04109 and T20493 were considered to be non-redundant clones. Redundancy is a common problem when dealing with the short error prone sequences of ESTs, however resistance genes frequently show tandem duplications, some of which are functional, and it is possible that the redundancy found within these clones is a reflection of this type of gene organization.

Forty-four of 49 R-ESTs were mapped in the N-fo-61-9 *B. napus* population. R-ESTs were mapped to every linkage group except N16 (no markers could be mapped to N16 due to monosomy in the 61-9 F1 plant; (Parkin *et al.* 1995), and ranged in number from

[a] GenBank accession number for *A. thaliana* ESTs.

[b] Estimated from the number of hybridizing fragments in EcoRI or EcoRV digested genomic *B. napus* DNA.

[c] Loci with multiple co-segregating RFLP bands indicated by '**'.

[d] Estimated from the number of hybridizing fragments from *A. thaliana* genomic DNA digested with ClaI, CfoI, DraI, EcoRI, EcoRV, HpaII, or ScaI.

[e] *A. thaliana* chromosomal map position as determined by Botella et al. (1997).

Table 2. Genomic locations and estimated copy number of loci detected with *A. thaliana* EST probes selected by homology to cloned disease resistance genes (R-ESTs).

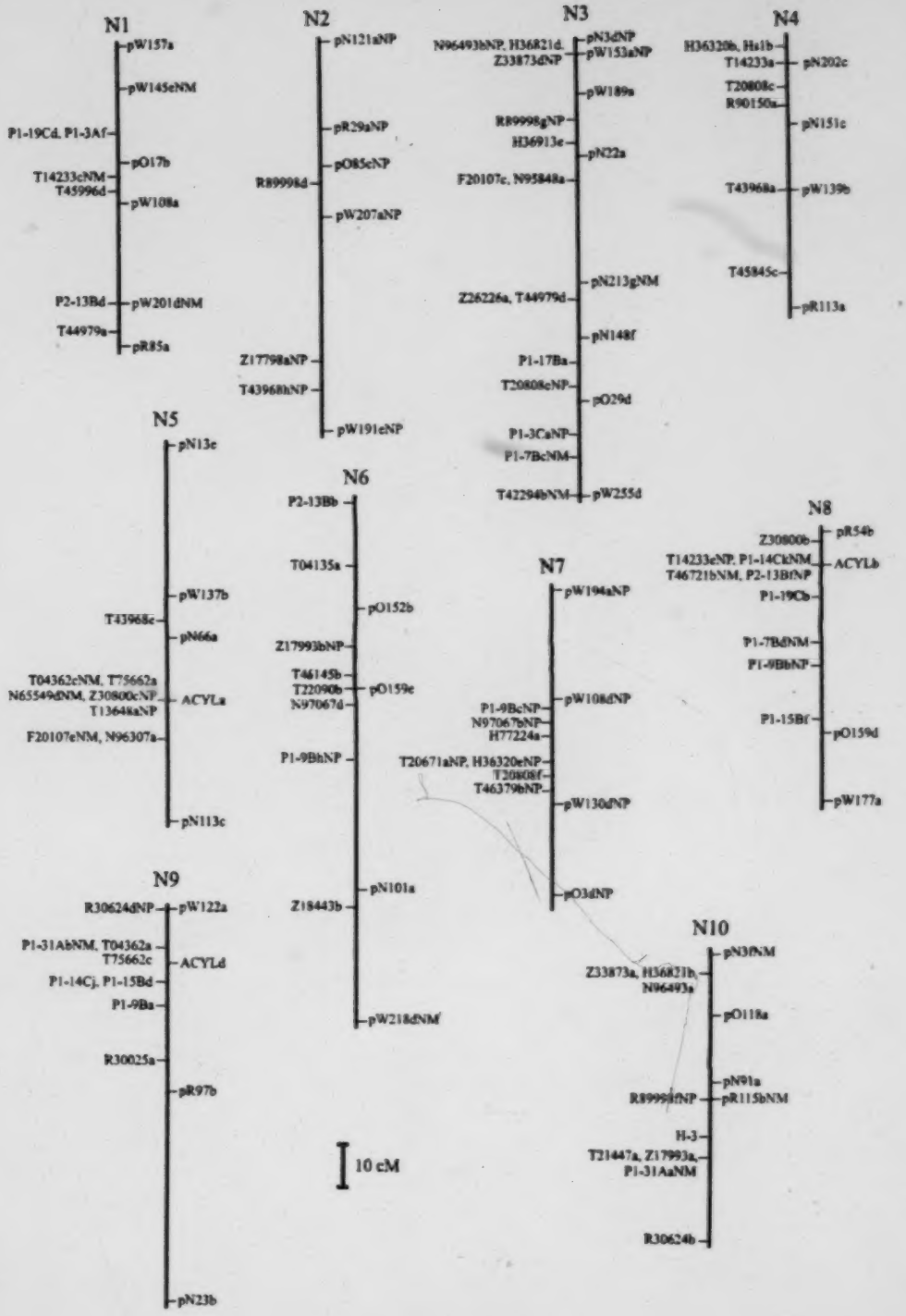
R-EST Probe [a]	Copy # *B.napus* [b]	Mapped *B.napus* Loci (# loci - Linkage groups)[c]	Copy # *A.thaliana* [d]	Mapped Loci in *A.thaliana* [c]	Map Pos. (Chr.) [e]
F20107/F20108	4	3-N13(a), N3(c), N5(eNM)	2	1-3(a)	2
H36320	6	3-N4(b), N7(eNP), N18(d)	1	1-3(a)	
H36821	6	4-N3(d), N13(aNP), N10(b)**, N19(c)**	2	1-5(c)	5
H36913	6	4-N12(aNP), N19(c), N13(b), N3(e)	1	no polymorphism	5
H37061	1	no polymorphism			
H77224	6	1-N7(a)	2	1-1(b)**	1
N65549	6	2-N11(bNP), N5(dNM)	3	2-1(b)**, 2(a)	
N65692		no polymorphism			
N95848	2	2-N3(a), N13(b)	2	1-1(a)	2
N96078	3-6	1-N13(aNP)	2	no polymorphism	4
N96307	2	2-N5(a), N15(b)	1	1-3(a)	
N96493	4-6	4-N10(a), N3(bNP), N19(d), N11(eNP)	1	1-5(a)	5
N96711	2	no polymorphism	2	1-5(a)	
N97067	6-8	3-N6(d), N7(bNP/eNP)	1-2	b(NM) unlinked	5
R29891		no polymorphism			
R30025	2	1-N9(a)	1	1-3(a)	3
R30624	7-9	3-N9(dNP), N10(bNP), N18(a)	2	2-4(b), 5(a)**	5
R64749		not R-EST			
R89998/T04109/T21150	6	5-N12(cNP), N3(gNP), N13(aNP), N10 (fNP), N19(bNM)	3	2-3(b), 5(a)	5
R90150	2	1-N4(a)	2	1-2(a)	1
T04135	4	2-N6(a), N15(b)	1	1-1(a)	1
T04362/T20493	4	2-N5(cNM), N9(a)	3	2-1(b)**, 2(a)	5
T13648	6-8	3-N12(bNM), N5(aNP), N18(c)	2	2-1(a)**, 5(b)**	
T14233	6	5-N1(cNM), N11(dNM), N14(bNP)**, N4(a)**, N8(eNP)	1	1-4(a)**	4
T20671	2	1-N7(aNP)	2	2-3(a), 4(b)	
T20808	6-8	5-N3(eNP), N4(c), N7(fNP), N19(aNM/dNP)	3	2-5(a/b)**	5
T21447	2	1-N10(a)	1	no polymorphism	1
T22090	3-4	2-N6(b)**, N15(a)	2	2-1(a), 4(b)	
T41629	2-4	1-N12(bNP)	1	1-1(a)	1
T41662	1	1-N17(a)	1	1-4(a)**	4
T42294	2-4	2-N3(b), N17(aNP)	1	1-1(a)**	
T43968	6	6-N2(hNP), N14(f/bNP), N5(e), N4(a), N19(c)	3	2-4(b), 5(a)**	5
T44979	4	3-N1(a), N3(d), N11(c)	1	no polymorphism	
T45845	4	3-N14(a/b), N4(c)	1-2	1-2(a)	2
T45996	5	3-N1(d), N19(c), N18(b)	1	1-3(a)	
T46145	2	2-N6(b), N15(a)**	1	1-1(a)	1
T46379	2	2-N14(c), N7(bNP)	1	1-2(a)	2
T46721	4	2-N13(a), N8(bNM)	1	1-4(a)	4
T75662	3	3-N5(a), N9(c), N15(d)**	1	1-1(a)	5
Z17798	4	2-N2(aNP), N12(cNP)	1	no polymorphism	5
Z17993	4-6	3-N6(bNP), N15(cNP), N10(aNM)	1	1-1(a)**	5
Z18443	4	2-N6(b), N17(a)	3	2-3(a), 1(dNM)	5
Z26226	4	2-N3(a), N13(b)	1	no polymorphism	5
Z30800	5-6	3-N13(dNM), N5(cNP), N8(b)	4	4-1(a/bNP/c), 4(dNP)	
Z33873	6	3-N3(dNP), N10(a), N19(b)	1	1-5(a)	5
Z34772	1	no polymorphism	2	1-1(a)	

three (N1, N2, N8, and N17) to 11 (N3) loci per linkage group. The loci detected by the R-ESTs were relatively evenly distributed over 18 linkage groups with apparent clustering only on N5, N7 and N8. In total, 103 loci detected by the R-ESTs were placed on the *B. napus* genetic linkage map (Figure 7).

In *A. thaliana*, 38 out of 46 R-ESTs were mapped to 48 loci positioned on all five chromosomes, with loci numbers ranging from five (chromosome 2) to 17 (chromosome 1) (Figure 8). Twelve loci were positioned in the major recognition complex (MRC) regions where multiple disease resistance loci are known to reside (Holub 1997), but overall, there was no apparent clustering of R-ESTs in these regions.

For most R-ESTs, the chromosomal location determined by Botella et al. (1997) was confirmed by this mapping; however for seven of the R-ESTs no loci were mapped to the previously determined linkage group. In the case of F20107, N95848, R90150, T04362, and Z18443, there were unmapped loci that could potentially correspond to the previously mapped position; but for each of T75662 and Z17993, only one locus was detected on a different linkage group than was determined by Botella et al. (1997) (Table 2). This was not unexpected, as it was acknowledged by Botella et al. (1997) that some map positions may be ambiguous due to high similarity among members of multigene families, or the presence of chimeric YAC clones.

Only one copy of Z17993 was detected in the *A. thaliana* genome, which was mapped to the top of chromosome 1 (mapped to chromosome 5 by Botella et al. 1997). Previously, Z17993 was found to co-segregate with *RPS5* on chromosome 1 in 97 RI lines (Warren *et al.* 1998) supporting our map position, and was used to clone the *RPS5*


N1
pW157a
pW145eNM
P1-19Cd, P1-3Af
pO17b
T14233cNM
T45996d
pW108a
P2-13Bd
pW201dNM
T44979a
pR85a
N2
pN121aNP
pR29aNP
R89998d
pO85cNP
pW207aNP
Z17798aNP
T43968hNP
pW191eNP
N3
N96493bNP, H36821d,
Z33873dNP
pN3dNP
pW153aNP
pW189a
R89998gNP
H36913e
pN22a
F20107c, N95848a
pN213gNM
Z26226a, T44979d
pN148f
P1-17Ba
T20808eNP
pO29d
P1-3CaNP
P1-7BcNM
T42294bNM
pW255d
N4
H36320b, Ha1b
T14233a
pN202c
T20808c
R90150a
pN151c
T43968a
pW139b
T45845c
pR113a
N5
pN13c
pW137b
T43968c
pN66a
T04362cNM, T75662a
N65549dNM, Z30800cNP,
T13648aNP
ACYLa
F20107eNM, N96307a
pN113c
N6
P2-13Bb
T04135a
pO152b
Z17993bNP
T46145b
T22090b
N97067d
pO159e
P1-9BbNP
pN101a
Z18443b
pW218dNM
N7
pW194aNP
pW108dNP
P1-9BcNP
N97067bNP
H77224a
T20671aNP, H36320eNP
T20808f
T46379bNP
pW130dNP
pO3dNP
N8
Z30800b
pR54b
T14233eNP, P1-14CkNM
T46721bNM, P2-13BfNP
ACYLb
P1-19Cb
P1-7BdNM
P1-9BbNP
P1-15Bf
pO159d
pW177a
N9
R30624dNP
pW122a
P1-31AbNM, T04362a
T75662c
ACYLd
P1-14Cj, P1-15Bd
P1-9Ba
R30025a
pR97b
pN23b
10 cM
N10
pN3fNM
Z33873a, H36821b,
N96493a
pO118a
pN91a
R89998fNP
pR115bNM
H-3
T21447a, Z17993a,
P1-31AaNM
R30624b

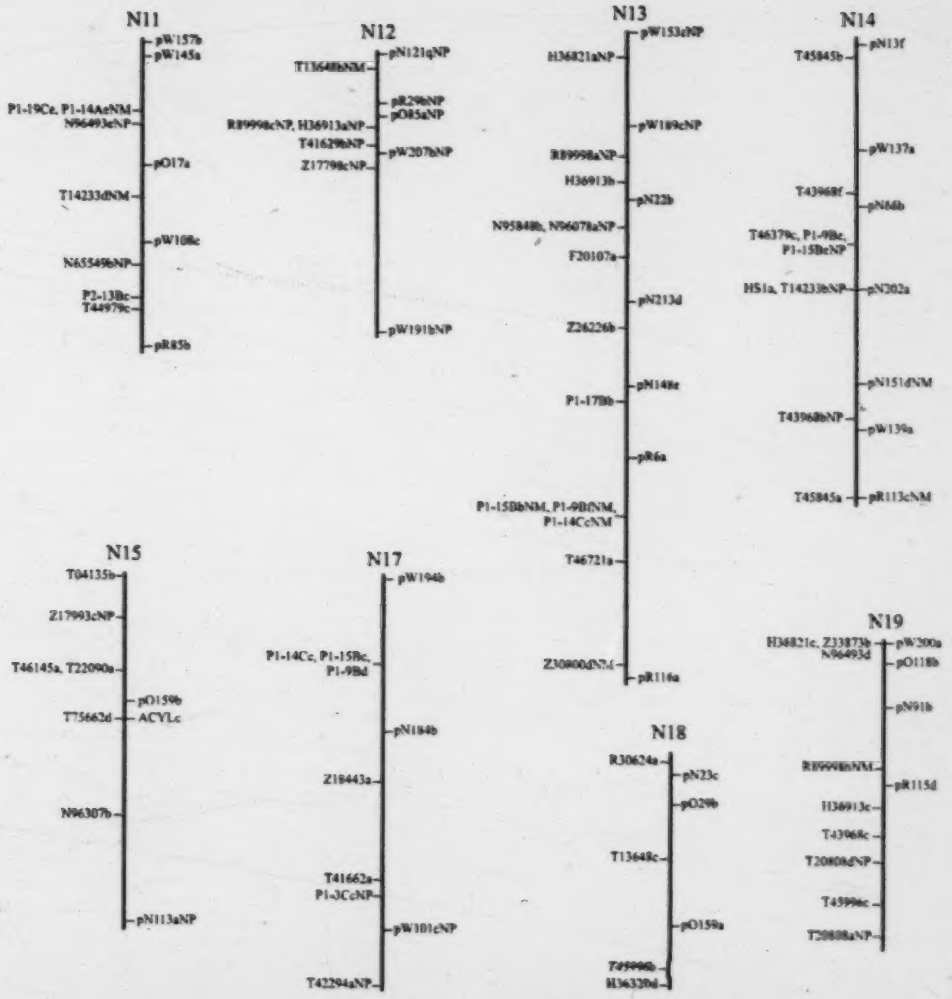
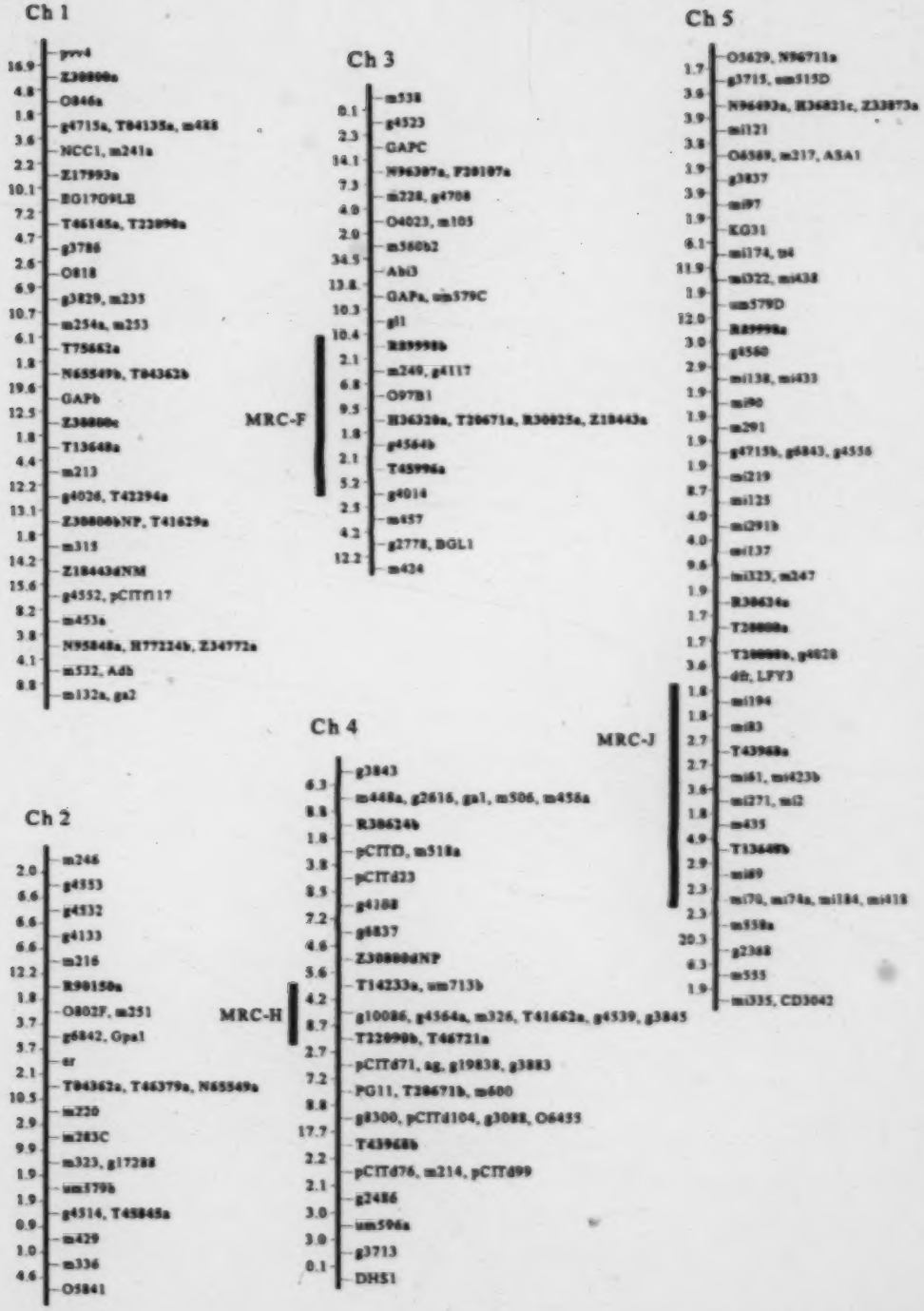


Figure 7. Genetic linkage map based on 30 doubled haploid (DH) lines of the N-fo-61-9 *B. napus* population, showing disease resistance gene-related sequences mapped in this study. Vertical bars represent linkage groups N1-N19, with selected framework markers from Parkin et al. (1995) on the right, and resistance-related markers on the left. Lower case letters following the locus designate multiple loci detected with the same RFLP probe. NP or NM after a locus name signify that only alleles from the N-o-9 or SYN1 parent respectively could be scored. R-ESTs are designated by GenBank accession numbers (eg. R89998), random *Brassica* R gene homologues by names beginning with 'P1-' or 'P2-' (eg. P1-17B), the *Brassica* Pto homologue by H-3, and the *Brassica* HS1^{pro-1} homologue by HS1. Map distances (Kosambi) between markers are drawn to scale.

Figure 8. Genetic linkage map of *A. thaliana* based on segregation in 30 recombinant (RI) lines of the Lister and Dean (1993) Columbia x Landsberg mapping population. The 5 chromosomes are represented by vertical bars, with markers on the right and map distances (Kosambi) in cM indicated on the left. R-ESTs are designated by GenBank accession number printed in bold face. Major recognition complex regions (MRC) (Holub, 1997) have been found to contain multiple disease resistance loci and are indicated by vertical bars to the left of chromosomes 3, 4, and 5.


Ch 1
16.9 pvv4
4.8 Z30800a
1.8 O846a
3.6 g4715a, T04135a, m488
2.2 NCC1, m241a
10.1 Z17993a
7.2 EG1709LE
4.7 T46148a, T23090a
2.6 g3786
6.9 O818
10.7 g3829, m235
6.1 m254a, m253
1.8 T75662a
19.6 N65549b, T04362b
12.5 GAPb
1.8 Z30800c
4.4 T13648a
12.2 m213
13.1 g4026, T42294a
1.8 Z30800bNP, T41629a
14.2 m315
15.6 Z18443dNM
8.2 g4552, pCITn17
3.8 m453a
4.1 N95848a, H77224b, Z34772a
8.8 m532, Adb
m132a, ga2
Ch 3
8.1 m538
2.3 g4523
14.1 GAPC
7.3 N96307a, F20107a
4.0 m228, g4708
2.0 O4023, m105
34.5 m560b2
13.8 Ab3
10.3 GAPa, um579C
10.4 g11
2.1 R89998b
6.8 m249, g4117
9.5 O97B1
1.8 H36320a, T20671a, R30025a, Z18443a
2.1 g4564b
5.2 T45996a
2.5 g4014
4.2 m457
12.2 g2778, BGL1
m434
MRC-F
Ch 5
1.7 O3629, N96711a
3.6 g3715, um515D
3.9 N96493a, H36823c, Z33873a
3.8 mi121
1.9 O6589, m217, ASA1
3.9 g3837
1.9 mi97
6.1 KG31
11.9 mi174, tt4
1.9 mi322, mi438
12.0 um579D
3.0 R89998a
2.9 g4560
1.9 mi138, mi433
1.9 mi90
1.9 m291
1.9 g4715b, g6843, g4556
1.9 mi219
8.7 mi125
4.0 mi291b
4.0 mi137
9.6 mi323, m247
1.9 R30624a
1.7 T20800a
1.7 T30099b, g4028
3.6 dft, LFY3
1.8 mi194
1.8 mi83
2.7 T43968a
2.7 mi61, mi423b
3.6 mi271, mi2
1.8 m435
4.9 T13648b
2.9 mi89
2.3 mi70, mi74a, mi184, mi418
2.3 m538a
20.3 g2388
6.3 m555
1.9 mi335, CD3042
MRC-J
Ch 4
6.3 g3843
8.8 m448a, g2616, ga1, m506, m458a
1.8 R30624b
3.8 pCITD, m518a
8.5 pCITd23
7.2 g4188
4.6 g6837
5.6 Z30800dNP
4.2 T14233a, um713b
8.7 g10086, g4564a, m326, T41662a, g4539, g3845
2.7 T22090b, T46721a
7.2 pCITd71, ag, g19838, g3883
8.8 PG11, T20671b, m600
17.7 g8300, pCITd104, g3088, O6435
2.2 T43968b
2.1 pCITd76, m214, pCITd99
3.0 g2486
3.0 um596a
0.1 g3713
DH51
MRC-H
Ch 2
2.0 m248
6.6 g4553
6.6 g4532
6.6 g4133
12.2 m216
1.8 R90150a
3.7 O802F, m251
5.7 g6842, Gpa1
2.1 er
10.5 T04362a, T46379a, N65549a
2.9 m220
9.9 m283C
1.9 m323, g17288
1.9 um579b
0.9 g4514, T45845a
1.0 m429
4.6 m336
O5841

gene from a BAC library. The hybridization pattern of Z17993 suggests it is present as a tandem repeat, which is consistent with the genomic organization of the *RPS5* locus (Aarts *et al.* 1998a). It is encouraging that an R-EST was used successfully to clone a functional R gene, thus demonstrating the plausibility of locating as yet unidentified R genes coincident with other R-EST loci.

Based on the minimum number of bands identified with each enzyme, the R-ESTs had estimated copy numbers ranging from one to more than nine in *B. napus*, and from one to four in *A. thaliana* (Table 2). Probes detecting multiple loci in *A. thaliana* tended to have higher copy number in *B. napus*. Many probes detected loci which could not be mapped due to lack of polymorphism between the parental lines or difficulty in resolving the banding patterns on autoradiographs.

3.1.2 Duplication of R-EST Loci in A. thaliana

The majority of R-ESTs in *A. thaliana* mapped to more that one location, in contrast to the findings of Botella et al. (1997), where only one locus was mapped per R-EST probe (Table 2). This latter discrepancy is the result of Botella et al. (1997) mapping many of the probes by PCR based screening of the *Arabidopsis* CIC YAC library. Many duplicate loci which can be detected by Southern hybridization would not be observed by amplification from primers specific to individual EST probes. An additional 17 loci in *A. thaliana* were mapped corresponding to 10 previously unmapped R-EST clones, using RFLP mapping in a recombinant inbred population (Lister and Dean 1993). It was apparent from our study that a number of the R-EST clones were present in multiple

copies in the *Arabidopsis* genome and in addition, 12 duplicate loci were positioned for 11 previously mapped R-ESTs.

As previously observed in *Arabidopsis*, duplicate loci were found on separate chromosomes, or distantly spaced on the same chromosome (McGrath *et al.* 1993) with the exception of T20808, which detected two loci (T20808a and T20808b) on *A. thaliana* chromosome 5 separated by only one recombination event. Of the five *B. napus* loci which could be positioned using T20808, two were positioned on N19. It is possible that these two loci are members of a tandemly repeated gene family in *A. thaliana* (McGrath *et al.* 1993) and that chromosomal rearrangement of a common ancestral chromosome has further separated these two loci in the *B. napus* genome.

3.1.3 Loci Organized as Tandem Repeats

As previously noted by Botella et al. (1997) a number of R-EST loci in *A. thaliana* (12 out of 52 total loci: 23%) appear to be present as tandem repeated sequences, based on the presence of multiple co-segregating bands (Table 2) (Figure 9). Multiple tandem repeated sequences have been observed in many well characterized disease resistance loci such as *Cf4* and *Cf9* from tomato, *M* and *N* from flax, and *Rp1* from maize. It has been proposed that tandem gene repeats, caused by unequal crossing over between sister chromatids or homologous chromosomes in meiosis (Ohno 1970) is a mechanism by which new disease resistance specificities evolve (Bennetzen and Hulbert 1992; Song *et al.* 1997). Although only two restriction enzymes were used to analysis the *B. napus* genomic DNA, very few of the detected loci (7 out of 103 total loci: 7%) showed co-segregating bands.

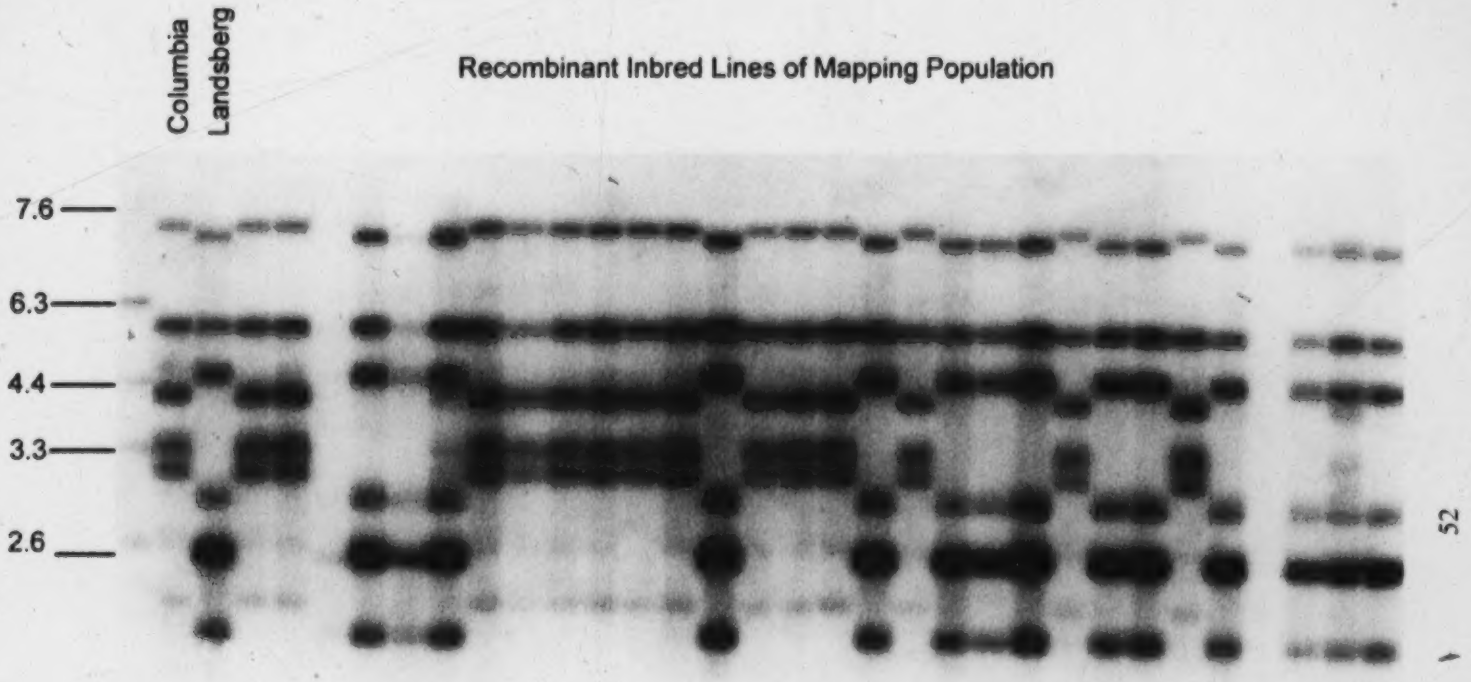


Figure 9. Autoradiograph of *A. thaliana* genomic DNA digested with DraI and probed with R-EST T14233. This banding pattern is characteristic of loci organized as tandem repeats. Multiple co-segregating bands can be seen for both the N-o-9 and N-f-2 alleles at this locus. Each member of the mapping population shows all bands from either parent. The most likely explanation for a single probe hybridizing to multiple restriction fragments of various sizes in this way is the presence of the target sequence organized as multiple linked copies. Numbers to the left are size standards in kb.

Interestingly T41662, which Botella et al. (1997) indicated as a putative EST for the downy mildew disease resistance gene *RPP5*, only detected a single locus in *B. napus*, inherited from *B. oleracea*, with a single band detected in each parent, 9 and 15 kb in length. *RPP5* is known to be part of a resistance gene cluster in *A. thaliana*, with at least four linked homologues present in the Landsberg ecotype spanning a distance of 50 kb (Parker *et al.* 1997). The RFLP locus T41662a was found to map in the MRC-H region of *A. thaliana* chromosome 4 where *RPP5* is known to reside. It also showed multiple co-segregating bands consistent with a tightly linked multigene family, supporting T41662 as an EST potentially representing the *RPP5* gene. It would be anticipated that if a similar organization was found in *B. napus* at the *RPP5* locus, numerous co-segregating bands, or a single band of high molecular weight would be observed in the mapping population.

Recent work on the organization of the *A. thaliana RPM1* locus in *B.napus* has shown there are six genomic regions homologous to the *RPM1* region present in the *B. napus* genome, but in four of these regions, the *RPM1* gene copy has been deleted (Grant *et al.* 1998). It may be that the deletion of duplicate gene copies is a common occurrence in the evolution of the *Brassica* diploid genomes, and that certain genes place an unnecessary genetic burden on the survival of the plant in the absence of pathogen pressure. Only as more genomic regions are compared at the physical level between *A. thaliana* and *Brassica* species will it be possible to distinguish between deletion of a gene, or extensive divergence of the DNA nucleotide sequence.

3.2 Genetic Mapping of *Brassica* R Gene Homologues

Eleven cloned R gene-like sequences from *B. napus* detected 33 loci on 11 linkage groups, ranging from one (N7 and N10) to six (N8) loci per linkage group (Figure 7). Many of the clones isolated by Joyeux et al. (1999) produced identical banding patterns in southern hybridization, and were assumed to be from redundant sequence, but it could be that these sequences represent tandem copies at the same locus. Joyeux et al. (1999) only mapped 14 loci with six polymorphic RFLP markers in the Major x Stellar *B. napus* mapping population, whereas I was able to map 33 loci with 11 markers in the highly polymorphic N-fo-61-9 population (Table 3).

Table 3. Copy number and linkage group assignment of *Brassica* R gene homologous sequences.

RFLP Probe [a]	Copy# *B.napus* [b]	Mapped *B.napus* Loci (# loci - Linkage groups)
P1-3A	7	1- N1(f)
P1-3C	6	2- N3(aNP), N17(cNP)
P1-7B	7-10	2- N8(dNM), N3(cNM)
P1-9B	7-10	7- N9(a), N7(cNP), N8(bNP), N6(hNP), N13(fNM), N17(d), N14(e)
P1-14A	6	1- N11(eNM)
P1-14C	10	4- N9(j), N8(kNM), N13(cNM), N17(e)
P1-15B	8	5- N9(d), N8(f), N13(bNM), N17(c), N14(eNP)
P1-17B	4	2- N3(a), N13(b)
P1-19C	4	3- N1(d), N8(b), N11(e)
P1-31A	5	2- N9(bNM), N10(aNM)
P2-13B	6	4- N1(d), N8(fNP), N6(b), N11(c)
H-3	1	1- N10
HS1	2	2- N14(a), N4(b)

[a] 'P1-' and 'P2-' probes are cloned *B. napus* sequences amplified via PCR using primers designed from the *A. thaliana* RPS2 resistance gene (Joyeux et al. 1999); H-3 is a *Brassica Pto* homologue; and HS1 is a *Brassica HS1*$^{pro-1}$ homologue.

[b] Estimated from the number of hybridizing fragments in EcoRI or EcoRV digested genomic DNA.

Copy numbers of these *Brassica* R gene-like sequences were generally high, ranging from four to 10. However, the *Brassica* homologue of the tomato *Pto* gene conferring

resistance to *Pseudomonas syringae* (H-3), and the *Brassica* homologue of the sugar beet nematode resistance gene, *HS1*$^{pro-1}$ (HS1) detected only one and two loci respectively.

Only three *Brassica* R gene homologue loci were common between the Major x Stellar, and N-fo-1-9 genetic maps. Both show a P1-3A locus on N1 (LG3), a P2-13B locus on N11 (LG 10+13), and a P1-15B locus on N9 (LG20). Judging from the high copy number of the *Brassica* homologues, many loci went unmapped in both this mapping study, and that of Joyeux et al. (1999), making alignment of the two maps unreliable.

One *Brassica* homologue amplified and cloned by Joyeux et al. (1999) (P1-16A) had 90% identity to *RPS2* at the nucleotide level. Unfortunately, this *RPS2* homologue could not be mapped in either the Major x Stellar or N-fo-61-9 *B. napus* populations. This probe detected only a single monomorphic locus in the N-fo-61-9 genomic DNA digested with either EcoRI or EcoRV. An attempt was made to map this interesting clone by conversion to a cleavable amplified polymorphic sequence (CAPS) PCR marker. With the aid of the OLIGO v.4.0 computer program, PCR primers were synthesized to amplify the P1-16A sequence from both the N-o-9 and N-f-2 parents (P1-16Afor - AGCTGATC ACAAAAGGTCATCAG; P1-16Arev - GATTCGGGCATATGGGATTT). The primers amplified a single product in both parents. Digestion of these PCR products with AluI, EcoRI, EcoRV, HinfIII, and Sau3AI failed to generate any polymorphisms between the parents. The fact that there is a single copy of this sequence in *B. napus*, and its apparent degree of sequence conservation make it interesting for future investigation.

On six of the *B. napus* linkage groups (N4, N8, N9, N10, N11 and N14), *Brassica* R gene homologues mapped coincident with *Arabidopsis* R-ESTs. It is possible that in

some of these cases, the *Brassica* homologue of the *Arabidopsis* gene represented by an R-EST has been amplified and cloned via degenerate PCR. It will be interesting to see if any of these pairs correspond to disease resistance function in both species.

There were three R-ESTs which showed a single copy in *B. napus*, H37061, T41662, and Z34772. It is interesting to note that two of these (H37061 and Z34772) were also not polymorphic with either EcoRI or EcoRV. If T41662 is considered to be an *RPP5* homologue (Botella *et al.* 1997) and P1-16A to be an *RPS2* homologue, a pattern of genomic organization in *B. napus* is suggested for sequences with high levels of homology to known disease resistance genes, including H-3 (*Pto* homologue) and HS1 ($HS1^{Pro-1}$ homologue). All these sequences are present in low copy number in *B. napus*, and may be relatively conserved in the case of T41662 and P1-16A, inferred from lack of polymorphism. The tandem repeat organization evident for T41662 in *A. thaliana* is not present in *B. napus*, supporting the speculation of deletion of superfluous R gene loci in *B. napus* (Grant *et al.* 1998; Henk *et al.* 1999). Perhaps interactions between homoeologous copies of R gene loci in the highly duplicated genome of *B. napus* are deleterious to the fitness of the plant, leading to selection pressure to maintain only one or two gene copies. These ideas are highly speculative based on the data collected, and further work will need to be done to answer these types of questions with greater certainty.

3.3 Collinearity Between *Arabidopsis* and *Brassica* Genomes Detected by R-ESTs

As expected, due to the dispersed nature of the mapped R-EST loci considered in this

study only limited regions of collinearity could be detected with the 37 markers successfully mapped in both the *A. thaliana* and *B. napus* genomes. R-EST loci mapped on *A. thaliana* chromosomes 1, 3, 4, and 5 detected collinear regions on multiple *B. napus* linkage groups; consistent with the hypothesized evolution of Brassicas by duplication and rearrangement of an ancestral genome similar in size and content to that of *A. thaliana* (Lagercrantz 1998). The patterns of collinearity uncovered for chromosomes 3, 4 and 5 were confined to previously defined blocks of synteny described in Scheffler et al. (1997), Cavell et al. (1998) and Osborn et al. (1997) respectively.

Scheffler et al. (1997) examined the copy number and map locations of members of a fatty-acid desaturase gene family in *B. napus*, and a collinear segment of *A. thaliana* chromosome 3. Six regions of the *B. napus* genome on homoeologous chromosomes N1 and N11, N3 and N13, and N5 and N15 were found to be collinear to a 30 cM segment of *A. thaliana* chromosome 3 containing 3 fatty-acid desaturase genes homologous to those in *B. napus*. R-EST probes N96307 and F20107 mapped in this 30 cM region of *A. thaliana*, and detected homologues on N3/N13 as well as N5/N15 (Figure 10). No loci were detected on linkage groups N1 or N11, as was found by Scheffler et al. (1997). This may be due to deletion events or sequence divergence preventing hybridization.

Collinearity relationships of *A. thaliana* chromosome 4 appear to be more complex, as it was found by Osborn et al. (1997) that the top arm contains only short conserved segments between the species, and no long stretches in this region appear to have homology to any single region or homoeologous chromosome group in *B. napus*. There are however, known collinear relationships with chromosome 4 which were confirmed by

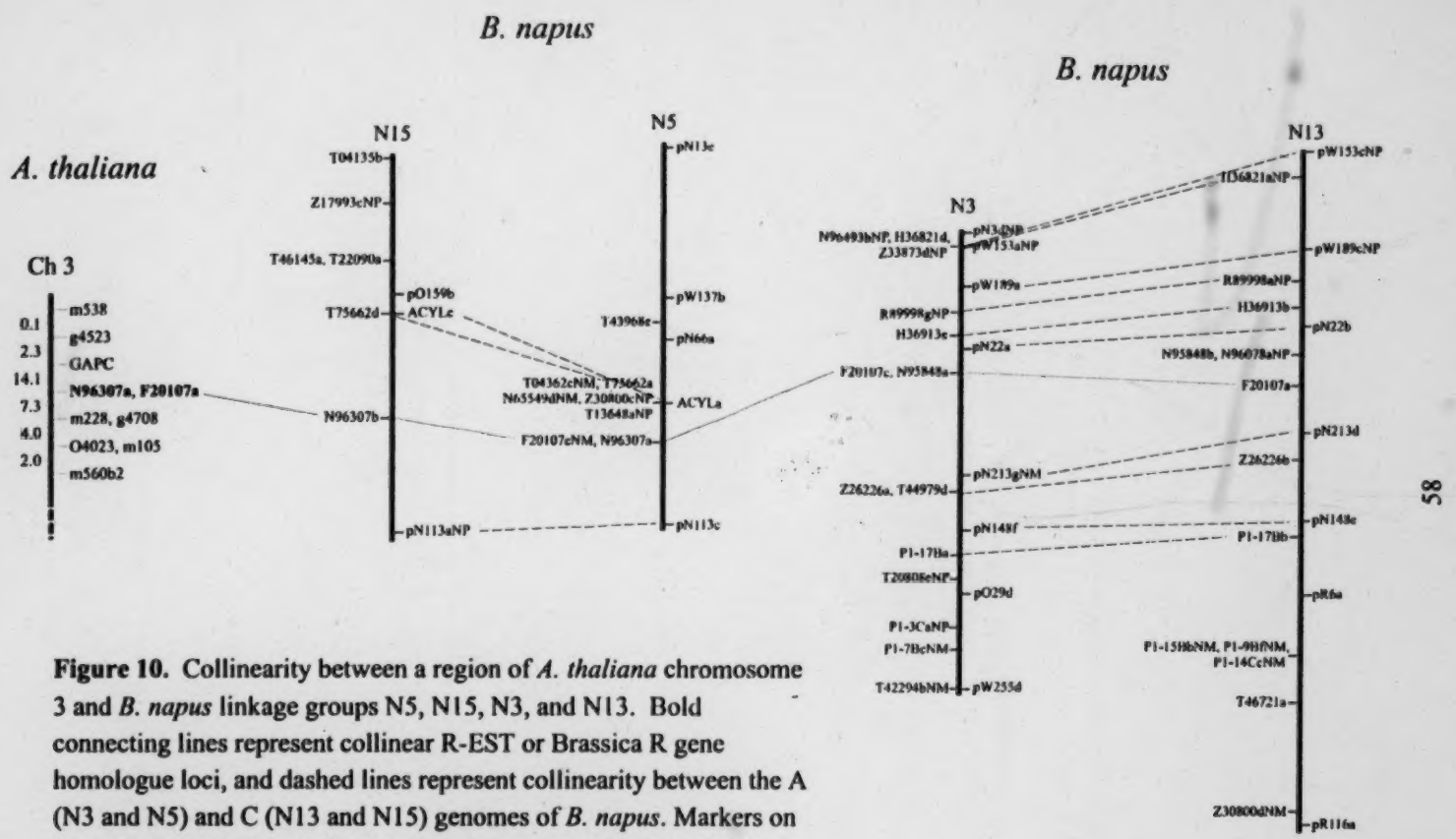


Figure 10. Collinearity between a region of *A. thaliana* chromosome 3 and *B. napus* linkage groups N5, N15, N3, and N13. Bold connecting lines represent collinear R-EST or Brassica R gene homologue loci, and dashed lines represent collinearity between the A (N3 and N5) and C (N13 and N15) genomes of *B. napus*. Markers on the right of the *B. napus* linkage groups are those added by Parkin et al. (1995). The collinear positions of N96307 and F20107 loci corroborate the mapping data of Scheffler et al. (1997).

the R-EST mapping data. Collinear regions in *B. napus* were identified for a 7.5 Mbp segment of *A. thaliana* chromosome 4 containing the *FCA* locus (Cavell *et al.* 1998). As with Sheffler et al. (1997), six collinear genomic regions were identified on pairs of linkage groups with previously determined homoeologous associations (N1/N11, N3/N17, N8/N18) (Cavell *et al.* 1998). R-ESTs T14233 and T41662 mapped to this region of *A. thaliana*, and homologues were detected in *B. napus* on linkage groups N1/N11, N4/N14, N17, and N8 (Figure 11). It is interesting that Cavell et al. (1998) did not detect the loci on N4/N14. This may have been the result of a lack of polymorphism for their markers in this region. Monomorphic loci will reduce estimates of the level of genomic duplication. For example, if 70% (p) of the detected loci are polymorphic, the probability of finding a polymorphism for both loci of a duplicate pair falls to 49% (p^2), and 34% (p^3) for a triplicated locus (Lagercrantz and Lydiate 1996). Scheffler et al. (1998) estimated that only 45% of the loci under investigation were polymorphic, so it is not unreasonable that a pair of homoeologous loci could have been missed.

Collinearity of flowering time genes in *B. napus* and *A. thaliana* was examined by Osborn et al. (1997). They mapped a cluster of markers from *A. thaliana* chromosome 5 around the flowering time loci *EMF1*, *FLC*, *FY*, and *CO* in *B. napus*, and linkage groups where at least two markers mapped included N2, N3, N10, N12, N13 and N19. One R-EST (R89998) mapped to this region of *A. thaliana* chromosome 5, and detected loci on *B. napus* linkage groups N3, N10, N12, N13, and N19, which agrees with the results of Osborn et al. (1997).

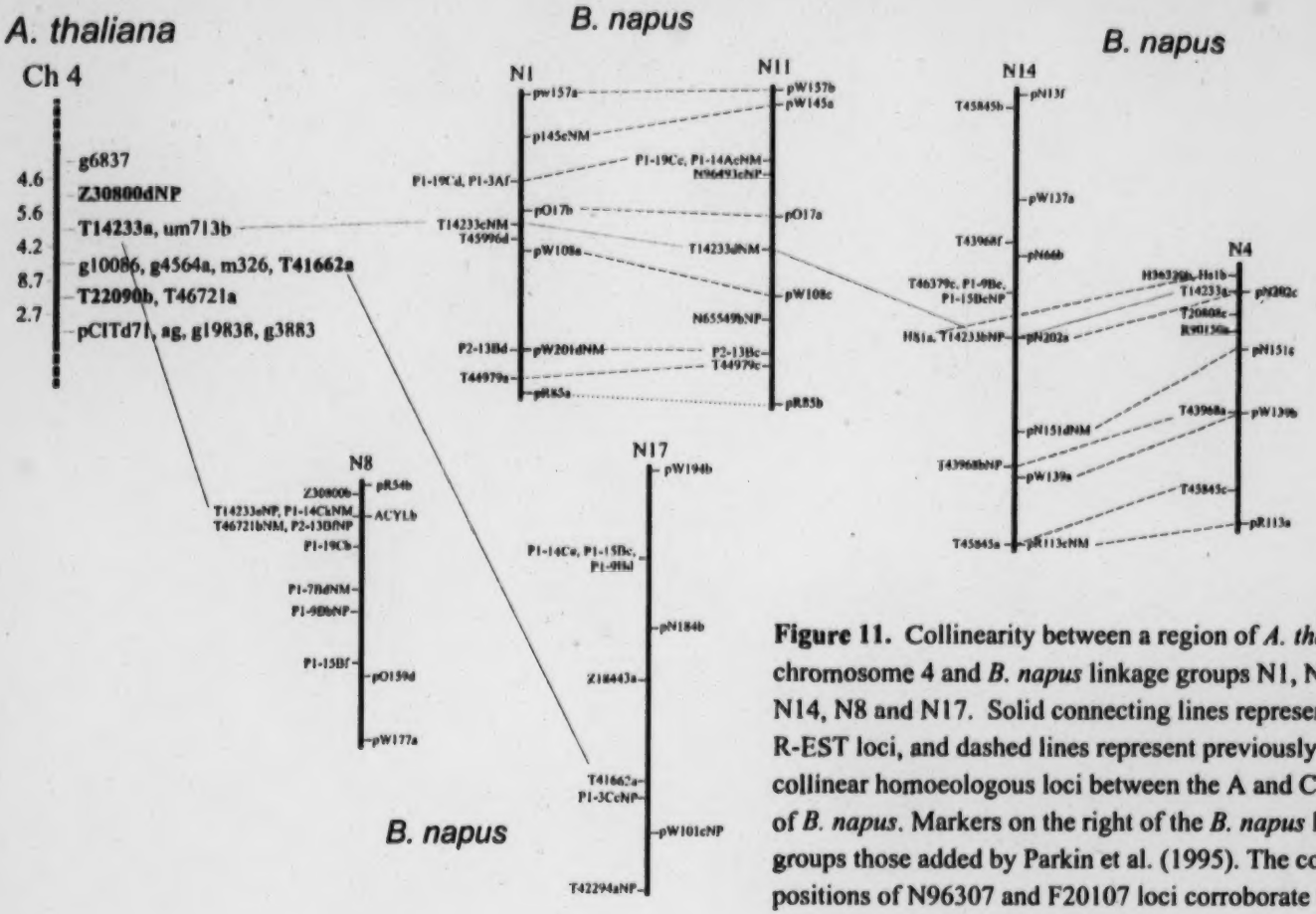


Figure 11. Collinearity between a region of *A. thaliana* chromosome 4 and *B. napus* linkage groups N1, N11, N4, N14, N8 and N17. Solid connecting lines represent collinear R-EST loci, and dashed lines represent previously mapped collinear homoeologous loci between the A and C genomes of *B. napus*. Markers on the right of the *B. napus* linkage groups those added by Parkin et al. (1995). The collinear positions of N96307 and F20107 loci corroborate the mapping data of Scheffler et al. (1997).

In this study, *A. thaliana* chromosome 1 had the highest number of mapped R-EST loci, making the identification of novel collinear relationships easier. The upper part of chromosome 1 was duplicated on N6 and N15, and a lower region of *Arabidopsis* chromosome 1 was duplicated on N5 and N15 (Figure 12). Distances between loci on collinear chromosomal stretches varied considerably. For example, the group of R-EST loci coincident on N5 (T04362cNM, T75662a, N65549dNM, Z30800cNP, and T13648aNP) are each separated by a number of recombination events on *A. thaliana* chromosome 1. Markers detecting loci on A. *thaliana* chromosome 1, and *B. napus* N6 and N15 show more conserved spacing. This region of collinearity between *A. thaliana* and *B. napus* shows a characteristic pattern which is as a result of the duplicated nature of the A (derived from *B. rapa*) genome and C (derived from *B. oleracea*) genome within *B. napus*. The top half of the C genome linkage group N15 is homoeologous with a region of the A genome linkage group N6, and the bottom half of the C genome linkage group N15 is homoeologous to a region on linkage group N5 of the A genome (Parkin *et al.* 1995) (Figure 13). Only four R-ESTs mapped to chromosome 2, and showed no clear collinearity to any specific regions of the *B. napus* genome.

In contrast to cereals, where it has been suggested that there has been rapid re-organization of resistance gene homologues over evolutionary time (Leister *et al.* 1998), no evidence of such a predilection in the evolution of *B. napus* and *A. thaliana* from a common Brassicaceae ancestor was detected by the R-EST mapping data. Duplicate loci detected by each of the R-EST probes mapped to homologous regions in the A and C genomes (Figure 13), and the limited regions of collinearity between *Arabidopsis* and

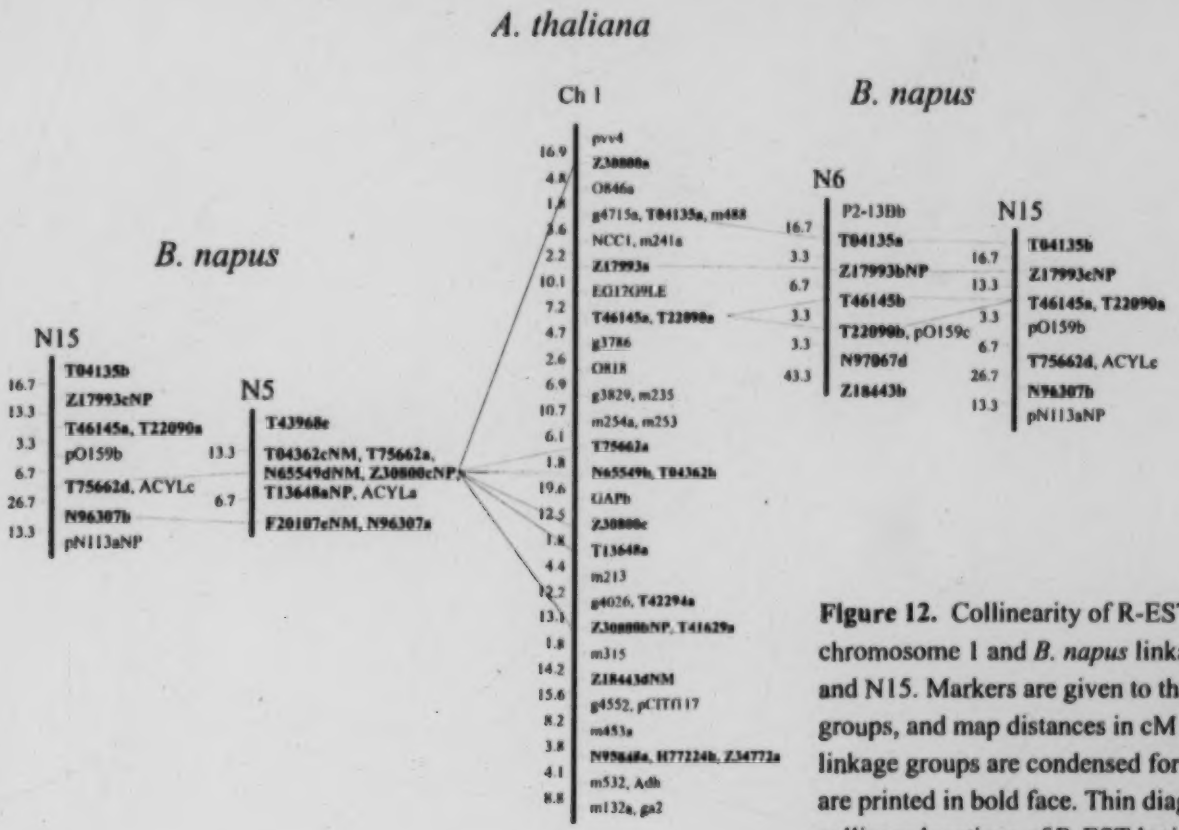


Figure 12. Collinearity of R-ESTs between *A. thaliana* chromosome 1 and *B. napus* linkage groups N5, N6, and N15. Markers are given to the right of the linkage groups, and map distances in cM to the left. *B. napus* linkage groups are condensed for clarity, and R-ESTs are printed in bold face. Thin diagonal lines denote collinear locations of R-EST loci between the two genomes. The top half of N15 is known to be homoeologous to a region on N6, and the lower half to a region on N5 (Parkin et al. 1995).

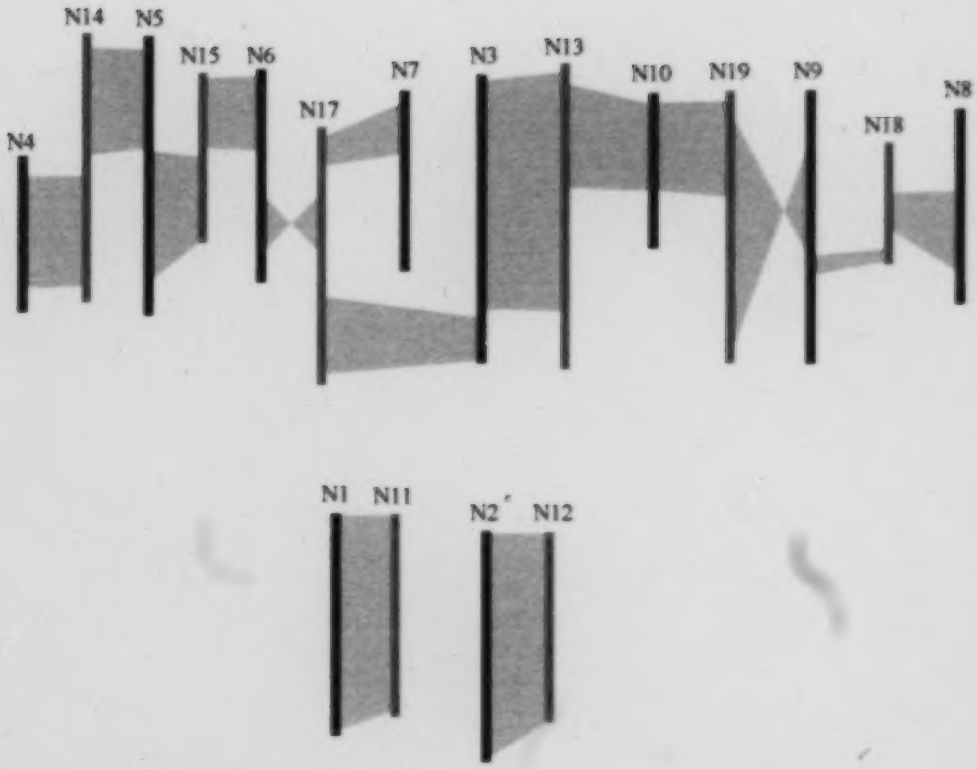


Figure 13. Schematic diagram of the homoeologous relationships among RFLP loci of the *B. napus* linkage groups. The Brassica A genome is represented by linkage groups N1-N10 (of *B. rapa* origin-black vertical bars), and the Brassica C genome is represented by linkage groups N11-N19 (of *B. oleracea* origin-grey vertical bars). R-EST loci mapping to multiple *B. napus* linkage groups followed this general pattern of organization.

Brassica which could be observed with such a dispersed set of probes showed conserved gene content and gene order. Perhaps the extensive duplication and rearrangement in *Brassica* genomes over evolutionary time is analogous to the re-organization of R genes within cereal genomes. *Brassica* genomes appear able to deal with chromosomal re-patterning very well, whereas the tight collinearity among the cereal genomes implies a negative selection pressure has kept chromosomes intact. Re-organization of R gene loci could be tolerated in cereals however, due to the increase in fitness afforded by multiple recognition specificities.

3.4 Conclusion

A number of researchers have shown that a subset of DNA sequences identified via homology to conserved R gene motifs are likely to represent disease resistance genes (Aarts *et al.* 1998a; Botella *et al.* 1997; Kanazin *et al.* 1996; Leister *et al.* 1996; Spielmeyer *et al.* 1998; Yu *et al.* 1996). This research will provide a useful resource for *Brassica* researchers who are attempting to clone such genes.

In this research, numerous candidate R genes have been mapped on every linkage group of *B. napus* and every chromosome of *A. thaliana*. Three *B. napus* regions were found to have slight clusters of loci, one of which is in the vicinity of two known blackleg resistance genes (on N7). Loci were also found in the MRC regions of *Arabidopsis*, which are known to contain multiple R genes, although there was not extensive clustering in these regions.

Twelve loci appear to be organized as tandem repeats in *A. thaliana*, and seven in *B.*

napus. This is a promising result, as this is a common organizational pattern seen in many of the cloned R genes from various plants. A much lower percentage (7%) of *B. napus* R gene homologue loci showed this type of tandem organization as compared to *A. thaliana* (23%). This may relate to the recent hypothesis that additional copies of R genes are deleterious (*B. napus* having higher copy numbers), with their deletion being under positive natural selection under certain conditions (Grant *et al.* 1998). More research must be done in this area however, before any definite conclusions can be drawn.

Arabidopsis R-ESTs mapped in both species outline homologous regions between the genomes. This allows the genomic sequence, markers, and physical libraries available in *Arabidopsis* (http://genome-www.stanford.edu/Arabidopsis/) to be used in *Brassica* research. Previously identified collinear regions were confirmed for *A. thaliana* chromosomes 3, 4, and 5, and novel collinearity was detected between *A. thaliana* chromosome 1 and *B. napus* linkage groups N5, N6 , and N15.

R gene homologues in the A and C genomes of *B. napus* were found to fit the known homoeologous linkage group relationships revealed by previous RFLP mapping. This would indicate that *Brassica* R gene homologues do not have a separate mechanism of evolution as has been proposed for R gene homologues in the grasses (Leister *et al.* 1998).

Only a small number of disease resistance loci have been placed on the *B. napus* map (Dion *et al.* 1995; Ferreira *et al.* 1995a; Ferreira *et al.* 1995b; Mayerhofer *et al.* 1997). In future research, these should be mapped in relation to the disease resistance markers presented in this research to identify candidate R genes to expedite their cloning,

characterization and utilization. This has already been done for the *B. napus LmR1* gene (Mayerhofer *et al.* 1997), and closely segregating R-ESTs have been identified. This will aid in the map-based cloning of this gene, with the ultimate goal of introducing durable blackleg resistance into future commercial *B. napus* cultivars.

IV. References

Aarts, M.G.M., Hekkert, B.T.L., Holub, E.B., Beynon, J.L., Stiekema, W.J., and Pereira, A. 1998a. Identification of R-gene homologous DNA fragments genetically linked to disease resistance loci in *Arabidopsis thaliana*. Mol. Plant-Microbe Interact. **11:** 251-258.

Aarts, N., Metz, M., Holub, E., Staskawicz, B.J., Daniels, M.J., and Parker, J.E. 1998b. Different requirements for *EDS1* and *NDR1* by disease resistance genes define at least two R gene-mediated signaling pathways in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **95:** 10306-10311.

Agrios, G.N. 1988. Plant pathology. Academic Press Inc., San Diego, California.

Ahn, S.N., Anderson, J.A., and Tanksley, S.D. 1993. Homoeologous relationships of rice, wheat and maize chromosomes. Mol. Gen. Genet. **241:** 483-490.

Ahn, S.N. and Tanksley, S.D. 1993. Comparative linkage maps of the rice and maize genomes. Proc. Natl. Acad. Sci. USA **90:** 7980-7984.

Allen, F.L. 1994. Usefulness of plant genome mapping to plant breeding. *In* Plant genome analysis. *Edited by* P.M. Gresshoff. CRC Press Inc., Boca Raton, FLA. pp. 11-18.

Anderson, P.A., Lawrence, G.J., Morrish, B.C., Ayliffe, M.A., Finnegan, E.J., and Ellis, J.G. 1997. Inactivation of the flax rust resistance gene *M* associated with loss of a repeated unit within the leucine-rich repeat coding region. Plant Cell **9:** 641-651.

Baker, B., Zambryski, P., Staskawicz, B., and Dinesh-Kumar, S.P. 1997. Signaling in plant-microbe interactions. Science **276:** 726-733.

Baker, C.J. and Orlandi, E.W. 1995. Active oxygen in plant pathogenesis. Annu. Rev. Phytopathol. **33:** 299-321.

Bennetzen, J.L. and Hulbert, S.H. 1992. Organization, instability and evolution of plant disease resistance genes. Plant Mol. Biol. **20:** 575-580.

Bent, A.F. 1996. Disease resistance genes: function meets structure. Plant Cell **8:** 1757-1771.

Bent, A.F., Kunkel, B.N., Dahlbeck, D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J., and Staskawicz, B.J. 1994. RPS2 of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science **265:** 1856-1860.

Boguski, M.S., Lowe, T.M.J., and Tolstoshev, S.H. 1993. db-EST- data base for 'expressed sequence tags'. Nature Genet. **4:** 332-333.

Bonerbiale, M.W., Plaisted, R.L., and Tanksley, S.D. 1988. RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics **120:** 1095-1103.

Botella, M.A., Coleman, M.J., Hughes, D.E., Nishimura, M.T., Jones, J.D.G., and Somerville, S.C. 1997. Map positions of 47 *Arabidopsis* sequences with sequence similarity to disease resistance genes. Plant J. **12:** 1197-1211.

Botella, M.A., Parker, J.E., Frost, L.N., Bittner-Eddy, P.D., Beynon, J.L., Daniels, M.J., Holub, E.B., and Jones, J.D.G. 1998. Three genes of the *Arabidopsis* RPP1 complex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. Plant Cell **10:** 1847-1860.

Boyes, D.C., Nam, J., and Dangl, J.L. 1998. The *Arabidopsis thaliana RPM1* disease resistance gene product is a peripheral plasma membrane protein that is degraded coincident with the hypersensitive response. Proc. Natl. Acad. Sci. USA **95:** 15849-15854.

Brears, T. and Ryals, J. 1994. Genetic engineering for disease resistance in plants. Agro-food-Industry Hi-Tech **July/Aug:** 10-13.

Cai, D., Kleine, M., Kifle, S., Harloff, H.-J., Sandal, N.N., Marcker, K.A., Klein-Lankhorst, R.M., Salentijn, E.M.J., Lange, W., Stiekema, W.J., Wyss, U., Grundler, M.W., and Jung, C. 1997. Positional cloning of a gene for nematode resistance in sugar beet. Science **275:** 832-834.

Caicedo, A.L., Schaal, B.A., and Kunkel, B.N. 1999. Diversity and molecular evolution of the RPS2 resistance gene in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **96:** 302-306.

Cavell, A.C., Lydiate, D.J., Parkin, I.A.P., Dean, C., and Trick, M. 1998. Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome. Genome **41:** 62-69.

Century, K.S., Shapiro, A.D., Repetti, P.P., Dahlbeck, D., Holub, E., and Staskawicz, B.J. 1997. *NDR1*, a pathogen-induced component required for *Arabidopsis* disease resistance. Science **278:** 1963-1965.

Chang, C., Bowman, J.L., DeJohn, A.W., Lander, E.S., and Meyerowitz, E.M. 1988. Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **85:** 6856-6860.

Chuong, P.V. and Beversdorf, W.D. 1985. High frequency embryogenesis through isolated microspore culture in *Brassica napus* and *Brassica carinata*. Plant Science **39:** 219-226.

Cooke, R., Raynal, M., Grellet, F., Delseny, M., Morris, P.-C., Guerrier, D., Giraudat, J., Quigley, F., Clabault, G., Li, Y.-F., Mache, R., Krivitzky, M., Gy, I.J.-J., Kreis, M., Lecharny, A., Parmentier, Y., Marbach, J., Fleck, J., Clément, B., Philipps, G., Hervé, C., Bardet, C., Tremousaygue, D., Lescure, B., Lacomme, C., Roby, D., Jourjon, M.-F., Chabriet, P., Charpenteau, J.-L., Desprez, T., Amselem, J., Chiapello, H., and Höfte, H. 1996. Further progress towards a catalogue of all *Arabidopsis* genes: analysis of a set of 5000 non-redundant ESTs. Plant J. **91:** 101-124.

Dangl, J.L. 1994. Enigmatic avirulence genes of phytopathogenic bacteria. *In* Bacterial pathogenesis of plants and animals: molecular and cellular mechanisms. Springer-Verlag, Berlin.

Dangl, J.L. 1995. Pièce de résistance: novel classes of plant disease resistance genes. Cell **80:** 363-366.

Dion, Y., Gugel, R.K., Rakow, G.F.W., Séguin-Swartz, G., and Landry, B.S. 1995. RFLP mapping of resistance to the blackleg disease [causal agent, *Leptosphaeria maculans* (Desm.) Ces. et de Not.] in canola (*Brassica napus* L.). Theor. Appl. Genet. **91:** 1190-1194.

Dixon, M.S., Hatzixanthis, K., Jones, D.A., Harrison, K., and Jones, J.D.G. 1998. The tomato *Cf-5* disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. Plant Cell **10:** 1915-1925.

Dixon, M.S., Jones, D.A., Keddie, J.S., Thomas, C.M., Harrison, K., and Jones, J.D.G. 1996. The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell **84:** 451-459.

Ellis, J.G., Lawrence, G.J., Luck, J.E., and Dodds, P.N. 1999. Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. Plant Cell **11:** 495-506.

Falk, A., Feys, B.J., Frost, L.N., Jones, J.D.G., Daniels, M.J., and Parker, J.E. 1999. *EDS1*, an essential component of R gene-mediated disease resistance in *Arabidopsis* has homology to eukaryotic lipases. Proc. Natl. Acad. Sci. USA **96:** 3292-3297.

Ferreira, M.E., Rimmer, S.R., Williams, P.H., and Osborn, T.C. 1995a. Mapping loci controlling *Brassica napus* resistance to *Leptosphaeria maculans* under different screening conditions. Phytopathology **85:** 213-217.

Ferreira, M.E., Williams, P.H., and Osborn, T.C. 1995b. Mapping of a locus controlling resistance to *Albugo candida* in *Brassica napus* using molecular markers. Phytopathology **85:** 218-220.

Fletcher, H.L. 1994. Possible loss of length conservation and reciprocity during recombination or conversion in tandem arrays. Genetics **138:** 511-518.

Flor, H.H. 1956. The complementary genic systems in flax and flax rust. Advances in Genetics **8:** 29-54.

Frye, C. and Innes, R.W. 1998. An *Arabidopsis* mutant with enhanced resistance to powdery mildew. Plant Cell **10:** 947-956.

Grant, M.R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R.W., and Dangl, J.L. 1995. Structure of the *Arabidopsis* RPM1 gene enabling dual specificity disease resistance. Science **269:** 843-846.

Grant, M.R., McDowell, J.M., Sharpe, A.G., Torres Zabala, M., Lydiate, D.J., and Dangl, J.L. 1998. Independent deletions of a pathogen-resistance gene in *Brassica* and *Arabidopsis* . Proc. Natl. Acad. Sci. USA **95:** 15843-15848.

Hammond-Kosack, K.E. and Jones, J.D.G. 1995. Plant disease resistance genes: unravelling how they work. Can. J. Bot. **73:** S495-S505.

Hammond-Kosack, K.E. and Jones, J.D.G. 1996. Resistance gene-dependent plant defense responses. Plant Cell **8:** 1773-1791.

Hammond-Kosack, K.E. and Jones, J.D.G. 1997. Plant disease resistance genes. Annu. Rev. Plant Physiol. Plant Mol. Biol. **48:** 575-607.

Henk, A., Warren, R.F., and Innes, R.W. 1999. A new Ac-like transposon of *Arabidopsis* is associated with a deletion of the RPS5 disease resistance gene. Genetics **151:** 1581-1589.

Holub, E.B. 1997. Organization of resistance genes in *Arabidopsis*. *In* The Gene-for-Gene Relationship in Plant-Parasite Interactions. *Edited by* I. Crute, E. Holub, and J. Burdon. CAB International, Wallingford, UK. pp. 5-26.

Johal, G.S. and Briggs, S.P. 1992. Reductase activity encoded by the *HM1* disease resistance gene in maize. Science **258:** 985-987.

Jones, D.A. and Jones, J.D.G. 1997. The role of leucine-rich repeat proteins in plant defense. Advances in Botanical Research **24:** 89-167.

Jones, D.A., Thomas, C.M., Hammond-Kosack, K.E., Balint-Kurti, P.J., and Jones, J.D.G. 1994. Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. Science **266:** 789-793.

Joyeux, A., Fortin, M.G., Mayerhofer, R., and Good, A.G. 1999. Genetic mapping of plant disease resistance gene homologues using a minimal *Brassica napus* L. population. Genome **42:** 735-743.

Kaminski, D., Rapp, B., and Celetti, M. 1997. Diseases of cereals, oilseeds and pulses. *In* Principles and practices of crop rotation. Saskatchewan Agriculture and Food, Saskatoon, SK. pp. 6-13.

Kanazin, V., Marek, L.F., and Shoemaker, R.C. 1996. Resistance gene analogs are conserved and clustered in soybean. Proc. Natl. Acad. Sci. USA **93:** 11746-11750.

Lagercrantz, U. 1998. Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics **150:** 1217-1228.

Lagercrantz, U. and Lydiate, D.J. 1996. Comparative genome mapping in *Brassica*. Genetics **144:** 1903-1910.

Lagercrantz, U., Putterill, J., Coupland, G., and Lydiate, D.J. 1996. Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. Plant J. **9:** 13-20.

Lagudah,E.S., Moullet,O., and Appels,R. 1997. Map-based cloning of a gene sequence encoding a nucleotide-binding domain and a leucine-rich region at the Cre3 nematode resistance locus of wheat. Genome **40:** 659-665.

Lander, E.S., Abrahamson, J., Barlow, A., Daley, M., Lincoln, S., and Newberg, L. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174-181.

Lawrence, G.J., Finnegan, E.J., Ayliffe, M.A., and Ellis, J.G. 1995. The *L6* gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene *RPS2* and the tobacco viral resistance gene *N*. Plant Cell **7:** 1195-1206.

Leister, D., Ballvora, A., Salamini, F., and Gebhardt, C. 1996. A PCR-Based approach for isolating pathogen resistance genes from potato with potential for wide applications in plants. Nature Genet. **14:** 421-428.

Leister, D., Kurth, J., Laurie, D.A., Yano, M., Sasaki, T., Devos, K., Graner, A., and Schulze-Lefert, P. 1998. Rapid reorganization of resistance gene homologues in cereal genomes. Proc. Natl. Acad. Sci. USA **95:** 370-375.

Levine, A., Tenhaken, R., Dixon, R.A., and Lamb, C. 1994. H_2O_2 from the oxidative burst orchestrates the plant hypersensitive disease resistance response. Cell **79:** 583-593.

Lister, C. and Dean, C. 1993. Recombinant inbred lines for mapping RFLP and phenotypic markers in *A. thaliana*. Plant J. **4:** 745-750.

Martin, G.B., Brommonschenkel, S.H., Chuwongse, J., Frary, A., Ganal, M.W., Spivey, R., Wu, T., Earle, E.D., and Tanksley, S.D. 1993. Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science **262:** 1432-1436.

Mayerhofer, R., Bansal, V.K., Thiagarajah, M.R., Stringam, G.R., and Good, A.G. 1997. Molecular mapping of resistance to *Leptosphaeria maculans* in Australian cultivars of *Brassica napus*. Genome **40:** 294-301.

McDowell, J.M., Dhandaydham, M., Long, T.A., Aarts, M.G.M., Goff, S., Holub, E.B., and Dangl, J.L. 1998. Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the RPP8 locus of Arabidopsis. Plant Cell **10:** 1861-1874.

McGrath, J.M., Jancso, M.M., and Pichersky, E. 1993. Duplicate sequences with a similarity to expressed genes in the genome of *Arabidopsis thaliana*. Theor. Appl. Genet. **86:** 880-888.

Mindrinos, M., Katagiri, F., Yu, G., and Ausubel, F.M. 1994. The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. Cell **78:** 1089-1099.

Nam, H.-G., Giraudat, J., den Boer, B., Moonan, F., Loos, W.B.D., Hauge, B.M., and Goodman, H.M. 1989. Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. Plant Cell **1:** 699-705.

Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S.C., Thomashow, M., Retzel, E., and Somerville, C. 1994. Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. Plant Physiology **106:** 1241-1255.

Ohmori, T., Murata, M., and Motoyoshi, F. 1998. Characterization of disease resistance gene-like sequences in near-isogenic lines of tomato. Theor. Appl. Genet. **96:** 331-338.

Ohno, S. 1970. Evolution by Gene Duplication. Springer-Verlag, New York. pp. 89-97.

Ori, N., Eshed, Y., Paran, I., Presting, G., Aviv, D., Tanksley, S.D., Zamir, D., and Fluhr, R. 1997. The *I2C* family from the wild disease resistance locus *I2* belongs to the nucleotide binding, Leucine-rich repeat superfamily of plant resistance genes. Plant Cell **9:** 521-532.

Osborn,T.C., Kole,C., Parkin,I.A.P., Sharpe,A.G., Kuiper,M., Lydiate,D.J., and Trick,M. 1997. Comparison of flowering time genes in *Brassica rapa, B. napus* and *Arabidopsis thaliana*. Genetics **146:** 1123-1129.

Parker, J.E., Coleman, M.J., Frost, L.N., Schmidt, R., van der Biezen, E.A., Moores, T., Dean, C., Daniels, M.J., and Jones, J.D.G. 1997. The *Arabidopsis* downey mildew resistance gene *RPP5* shares similarity to the toll and interleukin-1 receptors with *N* and *L6*. Plant Cell **9:** 879-894.

Parkin, I.A.P. and Lydiate, D.J. 1997. Conserved patterns of chromosome pairing and recombination in *Brassica napus* crosses. Genome **40:** 496-504.

Parkin, I.A.P., Sharpe, A.G., Keith, D.J., and Lydiate, D.J. 1995. Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). Genome **38:** 1122-1131.

Pawson, T. 1994. Introduction: protein kinases. FASEB J **8:** 1112-1113.

Prince, J.P., Pochard, E., and Tanksley, S.D. 1993. Construction of a molecular linkage map of pepper and a comparison of synteny with tomato. Genome **36:** 404-417.

Ryals, J.A., Neuenschwander, U.H., Willits, M.G., Molina, A., Steiner, H.-Y., and Hunt, M.D. 1996. Systemic acquired resistance. Plant Cell **8:** 1809-1819.

Sadowski, J., Gaubier, P., Delseny, M., and Quiros, C.F. 1996. Genetic and physical mapping in *Brassica* diploid species of a gene cluster defined in *Arabidopsis thaliana*. Mol. Gen. Genet. **251:** 298-306.

Sadowski, J. and Quiros, C.F. 1998. Organization of an *Arabidopsis thaliana* gene cluster on chromosome 4 including the RPS2 gene, in the *Brassica nigra* genome. Theor. Appl. Genet. **96:** 468-474.

Salmeron, J.M., Oldroyd, G.E.D., Rommens, C.M.T., Scofield, S.R., Kim, H., Lavelle, D.T., Dahlbeck, D., and Staskawicz, B.J. 1996. Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. Cell **86:** 123-133.

Scheffler,J.A., Sharpe,A.G., Schmidt,H., Sperling,P., Parkin,I.A.P., Lydiate,D.J., and Heinz,E. 1997. Desaturase multigene families of *Brassica napus* arose through genome duplication. Theor. Appl. Genet. **94:** 583-591.

Sharpe, A.G., Parkin, I.A.P., Keith, D.J., and Lydiate, D.J. 1995. Frequent nonriciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). Genome **38:** 1112-1121.

Song, W.Y., Pi, L.Y., Wang, G.L., Gardner, J., Holsten, T., and Ronald, P.C. 1997. Evolution of the Rice *Xa21* disease resistance gene family. Plant Cell **9:** 1279-1287.

Song, W.Y., Wang, G.L., Chen, L.L., Kim, H., Pi, L.Y., Holsten, T., Gardner, J., Wang, B., Zhai, W., Zhou, J., Fauquet, C., and Ronald, P.C. 1995. A receptor kinase-like protein encoded by the rice disease resistance gene Xa21. Science **270:** 1804-1806.

Spielmeyer, W., Robertson, M., Collins, N., Leister, D., Schulze-Lefert, P., Seah, S., Moullet, O., and Lagudah, E.S. 1998. A superfamily of disease resistance gene analogs is located on all homoeologous chromosome groups of wheat (*Triticum aestivum*). Genome **41:** 782-788.

Staskawicz, B.J., Ausubel, F.M., Baker, B.J., Ellis, J.G., and Jones, J.D.G. 1995. Molecular genetics of plant disease resistance. Science **268:** 661-667.

Staskawicz, B.J., Dahlbeck, D., and Keen, N.T. 1984. Cloned avirulence gene of *Pseudomonas syringae* pv. glycinea determines race-specific incompatibility on *Glycine max* (L.) Merr. Proc. Natl. Acad. Sci. USA **81:** 6024-6028.

Strange, R.N. 1993. Plant disease control. Chapman & Hall, London.

Tang, X., Frederick, R.D., Zhou, J., Halterman, D.A., Jia, Y., and Martin, G.B. 1996. Initiation of plant disease resistance by physical interaction of *AvrPto* and *Pto* kinase. Science **274:** 2060-2062.

Tenhaken, R., Levine, A., Brisson, L.F., Dixon, R.A., and Lamb, C. 1995. Function of the oxidative burst in hypersensitive disease resistance. Proc. Natl. Acad. Sci. USA **92:** 4158-4163.

Thilmony, R.L., Chen, Z., Bressan, R.A., and Martin, G.B. 1995. Expression of the tomato *Pto* gene in tobacco enhances resistance to Pseudomonas syringae pv tabaci expressing AvrPto. The plant cell 7:1529-1536. Plant Cell **7:** 1529-1536.

Thomas, C.M., Jones, D.A., Parniske, M., Harrison, K., Balint-Kurti, P.J., Hatzixanthis, K., and Jones, J.D.G. 1997. Characterization of the tomato Cf-4 gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. Plant Cell **9:** 2209-2224.

U, N. 1935. Genome-analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. Japanese Journal of Botany **9:** 389-452.

Van Deynze, A.E., Nelson, J.C., Yglesias, E.S., Harrington, S.E., Braga, D.P., McCouch, S.R., and Sorrells, M.E. 1995. Comparative mapping in grasses. Wheat relationships. Mol. Gen. Genet. **248:** 744-754.

Wang, G.L., Ruan, D.L., Song, W.Y., Sideris, S., Chen, L.L., Pi, L.Y., Zhang, S., Zhang, Z., Fauquet, C., Gaut, B.S., Whalen, M.C., and Ronald, P.C. 1998. Xa21D encodes a receptor-like molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. Plant Cell **10:** 765-779.

Warren, R.F., Henk, A., Mowery, P., Holub, E.B., and Innes, R.W. 1998. A mutation within the leucine-rich repeat domain of the *Arabidopsis* disease resistance gene *RPS5* partially suppresses multiple bacterial and downy mildew resistance genes. Plant Cell **10:** 1439-1452.

Warren, R.F., Merritt, P.M., Holub, E.B., and Innes, R.W. 1999. Identification of three putative signal transduction genes involved in R gene-specified disease resistance in *Arabidopsis*. Genetics **152:** 401-412.

Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C., and Baker, B. 1994. The product of tobacco mosaic virus resistance gene *N*: similarity to Toll and the Interleukin-1 receptor. Cell **78:** 1101-1115.

Woolhouse, H.W. 1992. Promoting crop protection by genetic engineering and conventional plant breeding: problems and prospects. *In* Plant genetic manipulation for crop protection. *Edited by* A.M.R. Gatehouse, V. A. Hilder, and D. Boulter. CAB International, Wallingford, UK. pp. 249-255.

Yang, Y., Shah, J., and Klessig, D.F. 1997. Signal perception and transduction in plant defense responses. Genes and Development **11:** 1621-1639.

Yu, Y.G., Buss, G.R., and Saghai Maroof, M.A. 1996. Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. Proc. Natl. Acad. Sci. USA **93:** 11751-11756.